

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 1999, 14:29:32 ; Search time 1261.22 Seconds

(without alignments)
12184.436 Million cell updates/sec

Title: US-09-001-039a-46

Sequence: 1 CTCGAGCTAAGATATTTTA.....ATTGATGATTCGGCCGC 4832

Scoring table: IDENTITY_NUC

Searched: 679419 seqs, 1590154680 residues

Database:

GenEmbl: *
1: gb_dal: *
2: gb_dal: *
3: gb_dal: *
4: gb_dal: *
5: gb_dal: *
6: gb_dal: *
7: gb_dal: *
8: gb_dal: *
9: gb_dal: *
10: gb_dal: *
11: gb_dal: *
12: gb_dal: *
13: gb_dal: *
14: gb_dal: *
15: gb_dal: *
16: gb_dal: *
17: gb_dal: *
18: gb_dal: *
19: gb_dal: *
20: gb_dal: *
21: gb_dal: *
22: gb_dal: *
23: gb_dal: *
24: gb_dal: *
25: gb_dal: *
26: gb_dal: *
27: gb_dal: *
28: gb_dal: *
29: gb_dal: *
30: gb_dal: *
31: gb_dal: *
32: gb_dal: *
33: gb_dal: *
34: gb_dal: *
35: gb_dal: *
36: gb_dal: *
37: gb_dal: *
38: gb_dal: *
39: gb_dal: *
40: gb_dal: *
41: gb_dal: *
42: gb_dal: *

BEST AVAILABLE COPY

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No. Query Match Length DB ID Description

1 4570.8 94.6 9354 5 AR003585 AR003585 Sequence

2	4192.4	86.8	4278	5	108644	108644 Sequence 4
3	4190.8	86.7	4281	5	108643	108643 Sequence 3
4	4045.8	83.7	4351	5	108641	108641 Sequence 1
5	4045.8	83.7	4358	5	108642	108642 Sequence 2
6	2468.8	51.1	9009	5	AR003710	AR003710 Sequence 2
7	2468.8	51.1	9009	5	131901	131901 Sequence 3
8	2468.8	51.1	9009	5	163424	163424 Sequence 1
9	2468.8	51.1	8967	5	171409	171409 Sequence 1
10	2468.8	51.1	9009	9	HUMFVIT1	K011740 Human coagu
11	2468.8	51.1	9029	9	HUMFVIT1C	M14113 Human coagu
12	2468.8	51.1	8967	10	HSEVIT1R	X01179 Human mRNA
13	2467.2	51.1	8241	5	A05328	A05328 Synthetic F
14	2467.2	51.1	8241	5	A07042	A07042 Artificial
15	2363	48.9	8831	5	E00527	E00527 Complete CD
16	2353.4	48.7	7440	5	108345	108345 Sequence 2
17	2338.4	48.4	7440	5	108457	108457 Sequence 1
18	2304.2	47.7	7272	5	105404	105404 Sequence 3
19	2293.4	47.5	7056	5	127063	127063 Sequence 1
20	2238	46.3	6999	5	177105	177105 Sequence 1
21	1813.6	37.5	7032	3	AF016234	AF016234 Canis fam
22	1806.4	37.4	7493	5	AR003712	AR003712 Sequence
23	1806.4	37.4	7493	5	163427	163427 Sequence 7
24	1806.4	37.4	7145	3	AF049489	AF049489 Canis fam
25	1805.6	37.4	7145	3	SSU49517	U49517 Sus. scrofa
26	1754.8	36.3	6539	5	102047	102047 Sequence 1
27	1723.2	35.7	1993	5	102054	102054 Sequence 8
28	1603.2	33.2	1728	5	127064	127064 Sequence 3
29	1400.4	25.7	1623	5	127064	M90707 Homo sapien
30	1007.4	20.6	2493	9	HUMF8C	104400 Sequence 35
31	996.8	20.8	3852	5	104400	E00422 Human genom
32	996.2	20.6	3852	5	E00422	AR003711 Sequence
33	879.6	18.2	1130	5	131900	131900 Sequence 1
34	879.6	18.2	1130	5	163426	163426 Sequence 5
35	879.6	18.2	1130	5	E00526	E00526 Part of PES
36	675.4	14.0	955	5	E00526	M88648 H. sapiens f
37	537	11.1	2709	9	HSA131818	A131818 Homo sapi
38	537	11.1	5348	10	102049	102049 Sequence 3
39	423.8	8.8	1728	5	BOVFACV	M81440 Bos taurus
40	313.6	6.5	6910	3	BOVFACV2A	M81441 Bos taurus
41	313.6	6.5	6895	3	HUMFVA	M14335 Human coagu
42	303.4	6.3	6893	10	A46255	A46255 Sequence 1
43	301.8	6.2	6909	5	A63218	A63218 Sequence 1
44	301.8	6.2	6909	5	HUMFVA	M16967 Human coagu
45	301.8	6.2	6909	9	HUMFVA	

ALIGNMENTS

RESULT 1
LOCUS AR003585 9354 bp DNA
DEFINITION Sequence 2 from patent US 5744326.
ACCESSION AR003585
NID G3964844
VERSION AR003585.1 GI:3964844
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 9354)
AUTHORS Ill.C.R. and Bidlingmaier, S.
TITLE Use of viral C18-acting post-transcriptional regulatory sequences to increase expression of intronless genes containing near-consensus splice sites
JOURNAL Patent: US 5744326-A 2 28-APR-1998;
FEATURES Location/Qualifiers
BASE COUNT 2506 a 2239 c 2161 g 2448 t
ORIGIN

Query Match 94.68; Score 4570.8; DB 5; Length 9354;
Best Local Similarity 98.88; Pred. No. 0;
Matches 4643; Conservative 0; Mismatches 12; Indels 45; Gaps 2;

14 TATTTAGAGAAATTAACCTTTTGGCTCTCCAGTTGAACTTTTGAGCAATAG---T 70
Db 2904 TATTTAGAGAAATTAACCTTTTGGCTCTCCAGTTGAACTTTTGAGCAATAGGCCAC 2963

71 CATGCAAAATAGACTCTCCACCTGCTTTCTTGTGCTTTTGCATTCGCTTAGTGC 130
Db 2964 CATGTTTATAGACTCTCCACCTGCTTTCTTGTGCTTTTGCATTCGCTTAGTGC 3023

131 CACCAGAAATAGACTCTCCAGCTGAGTGAAGTCAATGAGGACTATATGCAAGTATCT 190
Db 3024 CACCAGAAATAGACTCTCCAGCTGAGTGAAGTCAATGAGGACTATATGCAAGTATCT 3083

191 CGGTGAGCTGCTGTGAGACGCAAGATTTCCTCTAGAGTGCAGAAATCTTTTCATTCAA 250
3084 CGGTGAGCTGCTGTGAGACGCAAGATTTCCTCTAGAGTGCAGAAATCTTTTCATTCAA 3143

251 CACCTAGCTGTGTACAAAAAGCTGTTTGTAGAAATTCAGGATCACCTTTTCAACAT 310
Db 3144 CACCTAGCTGTGTACAAAAAGCTGTTTGTAGAAATTCAGGATCACCTTTTCAACAT 3203

311 CGCTAAGCCAAAGGCCACCCGATGGGTGCTGCTAGAGTCCATCCATCCAGGCTGAGGTTTA 370
Db 3204 CGCTAAGCCAAAGGCCACCCGATGGGTGCTGCTAGAGTCCATCCATCCAGGCTGAGGTTTA 3263

371 TGTATAGTGTGCTTATACACTTAAAGACATGAGCTTCCATCCTGTGACAGTTCATGCTGT 430
Db 3264 TGTATAGTGTGCTTATACACTTAAAGACATGAGCTTCCATCCTGTGACAGTTCATGCTGT 3323

431 TGGTATATCTACTGAGAAAGCTTCTAGAGGAGCTGAATATGATGATGACAGCAGTAAAG 490
Db 3324 TGGTATATCTACTGAGAAAGCTTCTAGAGGAGCTGAATATGATGATGACAGCAGTAAAG 3383

491 GGAGAGAAAGATGATAAAGTCTTCCCTGGTGAAGCAGTACATATGTCGGAGGTCCCT 550
Db 3384 GGAGAGAAAGATGATAAAGTCTTCCCTGGTGAAGCAGTACATATGTCGGAGGTCCCT 3443

551 GAAAGAAATGCTCAATGGCTCTGACCACTGTGCTTACCTACTCATATCTTTTCTCA 610
Db 3444 GAAAGAAATGCTCAATGGCTCTGACCACTGTGCTTACCTACTCATATCTTTTCTCA 3503

611 TGTGGACCTGTGTAAGAGCTTGAATTCAGAGCTCATTTGAGCCCTCTCTGATATGAGA 670
3504 TGTGGACCTGTGTAAGAGCTTGAATTCAGAGCTCATTTGAGCCCTCTCTGATATGAGA 3563

671 AGGAGATCTGGCCAAAGAAAGACAGACAGCTTGCACAATTTATCTACTTTTGTCTGT 730
Db 3564 AGGAGATCTGGCCAAAGAAAGACAGACAGCTTGCACAATTTATCTACTTTTGTCTGT 3623

731 ATTTGATGAGGAAAGATTGGCACTCAGAAACAAAGAACTCCTTGATGACAGATAGGGA 790
Db 3624 ATTTGATGAGGAAAGATTGGCACTCAGAAACAAAGAACTCCTTGATGACAGATAGGGA 3683

791 TGTGATATCTGCTGGGCTGCTGCTTAAATGCACAGTCAATGTTATGTAACAGGTC 850
3684^o TGTGATATCTGCTGGGCTGCTGCTTAAATGCACAGTCAATGTTATGTAACAGGTC 3743

851 TGTGCAAGTGTGATTTGATGTCACAGAAATCAGTCTATTGGCATGTATGGAATGGG 910
Db 3744 TGTGCAAGTGTGATTTGATGTCACAGAAATCAGTCTATTGGCATGTATGGAATGGG 3803

911 CACCACTCTGTAAGTGCATCAATATTCCTGAAAGTCAACATTTCTTGTGAGAACCA 970
Db 3804 CACCACTCTGTAAGTGCATCAATATTCCTGAAAGTCAACATTTCTTGTGAGAACCA 3863

971 TCGCAGAGGCTGCTGGAATCTCGCCATATCTTCTTACTGCTCAACAGCTGTGAT 1030
Db 3864 TCGCAGAGGCTGCTGGAATCTCGCCATATCTTCTTACTGCTCAACAGCTGTGAT 3923

1031 GGACCTTGACAGTTTCTACTGTTTGTGATATCTTCCACCAAGATGATGGATGA 1090

Db 3924 GGACCTTGACAGTTTCTACTGTTTGTGATATCTTCCACCAAGATGATGGATGA 3983

1091 AGCTTATGCAAGATGAGAGCTGTCCAGAGGAGACCCCAACAGCAATGAAAAATATGA 1150
Db 3984 AGCTTATGCAAGATGAGAGCTGTCCAGAGGAGACCCCAACAGCAATGAAAAATATGA 4043

1151 AGAAGCGGAAGACTATGATGATGATCTTACTGATCTGTAATGATGATGCTGAGGTTTGA 1210
Db 4044 AGAAGCGGAAGACTATGATGATGATCTTACTGATCTGTAATGATGATGCTGAGGTTTGA 4103

1211 TGTATGCAACTCTCTCTCTTATCCAAATTTGCTCAGTGTCCAGAAAGACTCTTAAAC 1270
Db 4104 TGTATGCAACTCTCTCTCTTATCCAAATTTGCTCAGTGTCCAGAAAGACTCTTAAAC 4163

1271 TTGGGATCATTTGCTGCTGGAAGAGAGACTGGGACTATGCTCCCTTGTGCTGCTGC 1330
Db 4164 TTGGGATCATTTGCTGCTGGAAGAGAGACTGGGACTATGCTCCCTTGTGCTGCTGC 4223

1331 CCCCAGTACAGAAAGTTATAAAGTCAATATTTGAACAATGGCCCTCAGCGGATTGGTAA 1390
Db 4224 CCCCAGTACAGAAAGTTATAAAGTCAATATTTGAACAATGGCCCTCAGCGGATTGGTAA 4283

1391 GAAATGCAAAAAAGTCCGATTTATGCGATACACAGATGAACCTTTAAGACTGTGAAGC 1450
Db 4284 GAAATGCAAAAAAGTCCGATTTATGCGATACACAGATGAACCTTTAAGACTGTGAAGC 4343

1451 TTATGAGATGATGAGAAATCTTGGACCTTACTTTATGAGGGAAGTTGAGACACACT 1510
Db 4344 TTATGAGATGATGAGAAATCTTGGACCTTACTTTATGAGGGAAGTTGAGACACACT 4403

1511 GTTGATATATTTAAGAATCAAGCAAGACAGCATATACATCTACCTCAGGAATCAC 1570
Db 4404 GTTGATATATTTAAGAATCAAGCAAGACAGCATATACATCTACCTCAGGAATCAC 4463

1571 TGTATCGCTCCTTTGTATTAAGAGATTAACAAAGGTGTAACAACTTTGAAGATTT 1630
Db 4464 TGTATCGCTCCTTTGTATTAAGAGATTAACAAAGGTGTAACAACTTTGAAGATTT 4523

1631 TCCAAATCTGCCAGAGAAATATTCAAATATAAATGACAGTGAAGATGAGGACC 1690
Db 4524 TCCAAATCTGCCAGAGAAATATTCAAATATAAATGACAGTGAAGATGAGGACC 4583

1691 AACTTAATCAGATCTCGGTGCTGACCCGCTATTACTGTAGTTTCTGTAATATGAGAG 1750
Db 4584 AACTTAATCAGATCTCGGTGCTGACCCGCTATTACTGTAGTTTCTGTAATATGAGAG 4643

1751 AGATCTAGCTTCAGGACATTTGGCCCTCTCTCATCTGTCACAAAGAAATCTGATGACA 1810
Db 4644 AGATCTAGCTTCAGGACATTTGGCCCTCTCTCATCTGTCACAAAGAAATCTGATGACA 4703

1811 AAGAGAAACAGATATATGACAGCAAGAGAAATGTCATCTGTTTCTGATTTGATGA 1870
Db 4704 AAGAGAAACAGATATATGACAGCAAGAGAAATGTCATCTGTTTCTGATTTGATGA 4763

1871 GAACCGAAGCTGTGTAAGTCTCAGAGAAATATACAAAGCTTTCTCCCAATCCAGCTGAGT 1930
Db 4764 GAACCGAAGCTGTGTAAGTCTCAGAGAAATATACAAAGCTTTCTCCCAATCCAGCTGAGT 4823

1931 GCAGCTTGAAGATCCAGGTTCCAAAGCCTCACAATCATATGACAGCAATGATGCTATGT 1990
Db 4824 GCAGCTTGAAGATCCAGGTTCCAAAGCCTCACAATCATATGACAGCAATGATGCTATGT 4883

1991 TTTGATAGTTTGAAGTGTGAGTTTGTTCATGAGGTTGACATACTGTCATTTCAAG 2050
Db 4884 TTTGATAGTTTGAAGTGTGAGTTTGTTCATGAGGTTGACATACTGTCATTTCAAG 4943

2051 CATTTGACACAGACTGACTTCTTTCTGTCTTCTCTGATATACCTTCAACACAA 2110
Db 4944 CATTTGACACAGACTGACTTCTTTCTGTCTTCTCTGATATACCTTCAACACAA 5003

2111 AATGGTATGAGACACACTCACCCTATTCCTCAGAGAAAGCTGCTCATGTC 2170

Db 5004 AATGGTATGAGACACACTCACCCATTCACAGAGAACTGCTTCATGTC 5063
Qy 2171 GATGAAAAACCCAGTCTATGATTCGGGTCGCCAACACTGACCTTTCGAAACAGG 2230
Db 5064 GATGAAAAACCCAGTCTATGATTCGGGTCGCCAACACTGACCTTTCGAAACAGG 5123
Qy 2231 CATGACCCCTTACTGAAAGTTTCTAGTTGTGACAAAGAACTGGTATTTTTCGAGGA 2290
Db 5124 CATGACCCCTTACTGAAAGTTTCTAGTTGTGACAAAGAACTGGTATTTTTCGAGGA 5183
Qy 2291 CAGTTATGAAATATTTTCAGCATCTTCTAGTAAATAACATGCAATGGAACCAAG 2350
Db 5184 CAGTTATGAAATATTTTCAGCATCTTCTAGTAAATAACATGCAATGGAACCAAG 5243
Qy 2351 CTCTCCCGAAA-----CCGACC 2368
Db 5244 CTCTCCCGAAATTCAGAACCCCTAGCCTAGGCAAAAGCAATTTAATGCCACCACC 5303
Qy 2369 AGCTTGAAGCCATCAAGGGGAAATACCTGCTACTCTTCAGTCAATCAAGAGA 2428
Db 5304 AGCTTGAAGCCATCAAGGGGAAATACCTGCTACTCTTCAGTCAATCAAGAGA 5363
Qy 2429 AATGACTATGATGATACCATATCAGTTGAATGAAGAAGATTTGACATTTATGA 2488
Db 5364 AATGACTATGATGATACCATATCAGTTGAATGAAGAAGATTTGACATTTATGA 5423
Qy 2489 TGAGAGTGAATATGAGACCCCGAGCTTTCAAAAGAAAACAGACACTATTTTATGC 2548
Db 5424 TGAGAGTGAATATGAGACCCCGAGCTTTCAAAAGAAAACAGACACTATTTTATGC 5483
Qy 2549 TGAGTGAAGGCTCTGGGATTTAGGATGATGATGCTCCCACTGTTCTTAAGAAAG 2608
Db 5484 TGAGTGAAGGCTCTGGGATTTAGGATGATGATGCTCCCACTGTTCTTAAGAAAG 5543
Qy 2609 GGCCTCAGATGGCAGTGTCCCTCAGTTCAAGAAAGTTTTCAGGATTTACTGATG 2668
Db 5544 GGCCTCAGATGGCAGTGTCCCTCAGTTCAAGAAAGTTTTCAGGATTTACTGATG 5603
Qy 2669 CTCCTTACTGAGCCCTTATACCTGAGAACTAAATGAACTTTGGACTCTGGGCC 2728
Db 5604 CTCCTTACTGAGCCCTTATACCTGAGAACTAAATGAACTTTGGACTCTGGGCC 5663
Qy 2729 AATATAGAGCAGAGTGAAGATATATCATGTAATCTTCAAAATCAGCCCTCG 2788
Db 5664 AATATAGAGCAGAGTGAAGATATATCATGTAATCTTCAAAATCAGCCCTCG 5723
Qy 2789 TCCCTATCTTCTATCTAGCCTTATTTCTATGAGAAAGATCAGAGCAAGAGCA 2848
Db 5724 TCCCTATCTTCTATCTAGCCTTATTTCTATGAGAAAGATCAGAGCAAGAGCA 5783
Qy 2849 ACCTAGAAAAAATTGTCAAGCCTAATGAACCAAACTTACTTTGAAAAGTCAACA 2908
Db 5784 ACCTAGAAAAAATTGTCAAGCCTAATGAACCAAACTTACTTTGAAAAGTCAACA 5843
Qy 2909 TCATATGSCACCCATTAAGATGATTTGACGCAAAAGCTTGCTATTTCTGATGT 2968
Db 5844 TCATATGSCACCCATTAAGATGATTTGACGCAAAAGCTTGCTATTTCTGATGT 5903
Qy 2969 TGACCTGAAAAAGATGCACTCAGGCTGATTTGACCCCTTCTGCTGCAACTGA 3028
Db 5904 TGACCTGAAAAAGATGCACTCAGGCTGATTTGACCCCTTCTGCTGCAACTGA 5963
Qy 3029 CACACTGAACCTGCTATGGAGACAAGTACAGTACAGAAATTTGCTGTTTTTAC 3088
Db 5964 CACACTGAACCTGCTATGGAGACAAGTACAGTACAGAAATTTGCTGTTTTTAC 6023
Qy 3089 CATCTTTATAGAACCAAAAGCTGCTACTCTGAAAAATATGAAAAAGAACTGAGGC 3148
Db 6024 CATCTTTATAGAACCAAAAGCTGCTACTCTGAAAAATATGAAAAAGAACTGAGGC 6083
Qy 3149 TCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAAATATGCTTCATGCAAT 3208
Db 6084 TCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAAATATGCTTCATGCAAT 6143

Qy 3209 CAATGCTACATTAATGATATACACTACTGCTTGTATGATGCTCAGATCAAGATTCG 3268
Db 6144 CAATGCTACATTAATGATATACACTACTGCTTGTATGATGCTCAGATCAAGATTCG 6203
Qy 3269 ATGGATCTGCTCAGCATGGGACAGCAATGAACATCATCTATTCATTTCAAGGACA 3328
Db 6204 ATGGATCTGCTCAGCATGGGACAGCAATGAACATCATCTATTCATTTCAAGGACA 6263
Qy 3329 TGTGTTCACTGTAAGAAAAAGAGAGTATTAATGAGCAGTACATCTATTCAGG 3388
Db 6264 TGTGTTCACTGTAAGAAAAAGAGAGTATTAATGAGCAGTACATCTATTCAGG 6323
Qy 3389 TGTGTTGAGACAGTGAATGTTACATCCAAAGCTGGAATTTGGCGGTGAATGCT 3448
Db 6324 TGTGTTGAGACAGTGAATGTTACATCCAAAGCTGGAATTTGGCGGTGAATGCT 6383
Qy 3449 TATGCGAGCATCTACATGCTGGATGAGACACTTTTCTGCTGACAGCAATAGTG 3508
Db 6384 TATGCGAGCATCTACATGCTGGATGAGACACTTTTCTGCTGACAGCAATAGTG 6443
Qy 3509 TCAGACTCCCTGGGAAATGCTCTGACACATTAAGATTTGATAGCTTCAGG 3568
Db 6444 TCAGACTCCCTGGGAAATGCTCTGACACATTAAGATTTGATAGCTTCAGG 6503
Qy 3569 ACAATATGAGACAGTGGCCCCAAAGCTGGCAGACTTATTTCCGATCAATCAATGC 3628
Db 6504 ACAATATGAGACAGTGGCCCCAAAGCTGGCAGACTTATTTCCGATCAATCAATGC 6563
Qy 3629 CTGAGACCAAGAGACCCCTTTCTGGATCAAGTGTATGTTGGACCAATGATAT 3688
Db 6564 CTGAGACCAAGAGACCCCTTTCTGGATCAAGTGTATGTTGGACCAATGATAT 6623
Qy 3689 TCAGGCACTCAAGACCCAGGCTGCCGTCAGAAAGTTTCCAGCCTCATCTCAGAT 3748
Db 6624 TCAGGCACTCAAGACCCAGGCTGCCGTCAGAAAGTTTCCAGCCTCATCTCAGAT 6683
Qy 3749 TATCATATGATATGCTTGTATGGAGAGAGTGGCAGACTTATCGAGGAATTCACATG 3808
Db 6684 TATCATATGATATGCTTGTATGGAGAGAGTGGCAGACTTATCGAGGAATTCACATG 6743
Qy 3809 AACCTATGCTCTTCTTGGCAATGATGATTCATCTGGGATTAACCAATATTTAA 3868
Db 6744 AACCTATGCTCTTCTTGGCAATGATGATTCATCTGGGATTAACCAATATTTAA 6803
Qy 3869 CCGTCAATTAATGCTGATACATCCGTTTCACCACTATTTATACATTCGAGCAC 3928
Db 6804 CCGTCAATTAATGCTGATACATCCGTTTCACCACTATTTATACATTCGAGCAC 6863
Qy 3929 TCTTGCATGGAATGATGGCTGTGATTAATAGTTGACAGCATGCCATTTGGAAATGA 3988
Db 6864 TCTTGCATGGAATGATGGCTGTGATTAATAGTTGACAGCATGCCATTTGGAAATGA 6923
Qy 3989 GAGTAAGCAATATCAGATGACAGATTAATGCTTCACTTACTTACCATATGTTGC 4048
Db 6924 GAGTAAGCAATATCAGATGACAGATTAATGCTTCACTTACTTACCATATGTTGC 6983
Qy 4049 CACCTGCTCTCTTCAAAAGCTGCACTTCCACCAAGGAGAGATGCTGAGAGC 4108
Db 6984 CACCTGCTCTCTTCAAAAGCTGCACTTCCACCAAGGAGAGATGCTGAGAGC 7043
Qy 4109 TCAGTGAATATCCAAAGAGTGGCTCAAGTGAATTCAGAGACATGAAGTCAAC 4168
Db 7044 TCAGTGAATATCCAAAGAGTGGCTCAAGTGAATTCAGAGACATGAAGTCAAC 7103
Qy 4169 AGGATTAATCTACAGGAGATTAATCTGCTTACAGCATGTATGGAAGATTCCT 4228
Db 7104 AGGATTAATCTACAGGAGATTAATCTGCTTACAGCATGTATGGAAGATTCCT 7163
Qy 4229 CATCTCCAGACATCAAGATGCCATCAGTGAATCTTTTTCAGATGCAAGTAAA 4288
Db 7164 CATCTCCAGACATCAAGATGCCATCAGTGAATCTTTTTCAGATGCAAGTAAA 7223

QY 4289 GGTTCCTCAGGAAATCAAGACTCTCTACACCGCTGGTGAACCTCTCTAGACCCACCGTT 4348
|||||
Db 7224 GGTTCCTCAGGAAATCAAGACTCTCTACACCGCTGGTGAACCTCTCTAGACCCACCGTT 7283
QY 4349 ACTGACTCGTACCTTCGAAATTCACCCCGAGAGTTGGGTGCACACAGATTGCCGTGAGAT 4408
|||||
Db 7284 ACTGACTCGTACCTTCGAAATTCACCCCGAGAGTTGGGTGCACACAGATTGCCGTGAGAT 7343
QY 4409 GAGGTTCTGGGTGGGAGGACAGACCTCTCTAGAGGTGGCCACTGACAGACCTGCC 4468
7344 GAGGTTCTGGGTGGGAGGACAGACCTCTCTAGAGGTGGCCACTGACAGACCTGCC 7403
QY 4469 ACAGCGGTACCTCTCCCTCTCAGCTCCAGGAGAGTCCCTCCGCTGGCTCTCTA 4528
Db 7404 ACAGCGGTACCTCTCCCTCTCAGCTCCAGGAGAGTCCCTCCGCTGGCTCTCTA 7463
QY 4529 CCTTTGTCTAAATCTAGACAGACCTGCTTGAAGCTCTCTGAATTAATCATCATAGT 4588
7464 CCTTTGTCTAAATCTAGACAGACCTGCTTGAAGCTCTCTGAATTAATCATCATAGT 7523
QY 4589 CCTGCAATTTCTTTGGTGGGGGGCCAGAGGGTGCATCCAAATTAACCTTTACCTA 4648
Db 7524 CCTGCAATTTCTTTGGTGGGGGGCCAGAGGGTGCATCCAAATTAACCTTTACCTA 7583
QY 4649 TTTTCGACGTGCTCCAG 4668
Db 7584 TTTTCGACGTGCTCCAG 7603

RESULT 2
LOCUS 108644 4278 bp PAT 14-NOV-1994
DEFINITION Sequence 4 from Patent WO 8800831.
ACCESSION 108644
NID 9588650
VERSION 108644.1 GI:588650
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4278)
AUTHORS Pasek M.P.
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR
POLYPEPTIDES IN HIGH YIELDS
PATENT: WO 8800831-A 4 11-FEB-1988;
JOURNAL Location/Qualifiers
FEATURES
BASE COUNT 1244 a 942 c 946 g 1146 t
ORIGIN

Query Match 86.8%; Score 4192.4; DB 5; Length 4278;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 4271; Conservative 0; Mismatches 1; Indels 48; Gaps 3;

QY 129 GGCACAGAGATGACTGCTGGTGCAGTGAAGTCTCATGGAGCTATATGCAAGTAT 188
Db 1 GGCACAGAGATGACTGCTGGTGCAGTGAAGTCTCATGGAGCTATATGCAAGTAT 60
QY 189 CTGGGTGAGCTGCTGTGAGCAGCAAGATTCTCTAGAGTGGCAAAATCTTTCCATTG 248
Db 61 CTGGGTGAGCTGCTGTGAGCAGCAAGATTCTCTCTAGAGTGGCAAAATCTTTCCATTG 120
QY 249 AACACCTGAGTGTGTACAAAAAGACTGTGTTGTAGAAATCAAGGATCACTTTTCAAC 308
Db 121 AACACCTGAGTGTGTACAAAAAGACTGTGTTGTAGAAATCAAGGATCACTTTTCAAC 180
QY 309 ATCGGTAAAGCAGGCGACCTGGAGGGTCTGCTAGTCTTACCATTCAGGCTGAGGTT 368
Db 181 ATCGGTAAAGCAGGCGACCTGGAGGGTCTGCTAGTCTTACCATTCAGGCTGAGGTT 240

QY 369 TATGATACGTGTGATTAACCTTAAGAAACAGCTTCCATCTGTCAGTCTTATGCT 428
Db 241 TATGATACGTGTGATTAACCTTAAGAAACAGCTTCCATCTGTCAGTCTTATGCT 300
QY 429 GTTGGTGTATCTTACCTGGAAGCTTCTGAGGAGCTGAATATGATGATAGACCGATCA 488
Db 301 GTTGGTGTATCTTACCTGGAAGCTTCTGAGGAGCTGAATATGATGATAGACCGATCA 360
QY 489 AGGAGAGAAAGATGATTAAGTCTTCCCTGGTGAAGCCATATGATGTCGAGAGTC 548
Db 361 AGGAGAGAAAGATGATTAAGTCTTCCCTGGTGAAGCCATATGATGTCGAGAGTC 420
QY 549 CTGAAAGAAATGTCATTAAGTCCCTGCTGACCACTGTGCTTACCTACTCATATCTTCT 608
Db 421 CTGAAAGAAATGTCATTAAGTCCCTGCTGACCACTGTGCTTACCTACTCATATCTTCT 480
QY 609 CATGTGACCTGTGTAAGACTGTAATTCAGGCTCATTTGGAGCCCTCTACTATATGTA 668
Db 481 CATGTGACCTGTGTAAGACTGTAATTCAGGCTCATTTGGAGCCCTCTACTATATGTA 540
QY 669 GAAGGAGTCTGCGCAAGGAAAGACACAGACCTTGCACAAATTTATCTACTTTTCT 728
Db 541 GAAGGAGTCTGCGCAAGGAAAGACACAGACCTTGCACAAATTTATCTACTTTTCT 600
QY 729 GTATTTGATGAAGGAAAAAGTTGGCACTCAGAACAAAGAACTCTTGTATGACGATAG 788
Db 601 GTATTTGATGAAGGAAAAAGTTGGCACTCAGAACAAAGAACTCTTGTATGACGATAG 660
QY 789 GATGCTGATCTGCTCGGGGCTCGGCTAAATGCACACAGTCAATGTTATTAACAGG 848
Db 661 GATGCTGATCTGCTCGGGGCTCGGCTAAATGCACACAGTCAATGTTATTAACAGG 720
QY 849 TCTCTG--CCAGGTCTGATTTGATGCCACAGAAATCATGTTGATGATGATGGA 905
Db 721 TCTCTGATCACAGGTCTGATTTGATGCCACAGAAATCATGTTGATGATGATGGA 780
QY 906 ATGGGACACCTCTGAGTGCATCATATTTCTCGAAGGTCACACATTTCTTTGAG 965
Db 781 ATGGGACACCTCTGAGTGCATCATATTTCTCGAAGGTCACACATTTCTTTGAG 840
QY 966 AACATTCGACAGGCTCTCTGGAATTCGCAATATCTTCTCTACGTCACAACTTC 1025
Db 841 AACATTCGACAGGCTCTCTGGAATTCGCAATATCTTCTCTACGTCACAACTTC 900
QY 1026 TTGATGACCTTGGACAGTTTCTACTGTTTGTCTATATCTCTCCACCAACATGATGC 1085
Db 901 TTGATGACCTTGGACAGTTTCTACTGTTTGTCTATATCTCTCCACCAACATGATGC 960
QY 1086 ATGAGGCTTATGTCAAAGTGAAGCTGTCCAGAGGAAACCCCACTACGATGAAAT 1145
Db 961 ATGAGGCTTATGTCAAAGTGAAGCTGTCCAGAGGAAACCCCACTACGATGAAAT 1020
QY 1146 AATGAAGAAGCGAAGACTATGATGATGATCTACTGATTCGAATGATGATGTCAGG 1205
Db 1021 AATGAAGAAGCGAAGACTATGATGATGATCTACTGATTCGAATGATGATGTCAGG 1080
QY 1206 TTTGATGATGACACTCTCTCTCTTATCCAAATTCGCTCAGTTGCCAAGAAGCATCT 1265
Db 1081 TTTGATGATGACACTCTCTCTCTTATCCAAATTCGCTCAGTTGCCAAGAAGCATCT 1140
QY 1266 AAAACTTGGGTACATTAATGCTGCTGTAAGAGAGAGCTGGGCTATGCTCCCTATCT 1325
Db 1141 AAAACTTGGGTACATTAATGCTGCTGTAAGAGAGAGCTGGGCTATGCTCCCTATCT 1200
QY 1326 CTCGCCCGGATGACAGAGATTATTAAGTCAATATTTGAACAATGAGGCTCAGGAGTT 1385
Db 1201 CTCGCCCGGATGACAGAGATTATTAAGTCAATATTTGAACAATGAGGCTCAGGAGTT 1260
QY 1386 GGTAGGAAGTACAAAAAGTCCGATTTATGSCATPACAGATGAACCTTTAAGACTGT 1445
Db 1261 GGTAGGAAGTACAAAAAGTCCGATTTATGSCATPACAGATGAACCTTTAAGACTGT 1320
QY 1446 GAAGCTATTCAGCATGATCAGGATCTTGGGACCTTACTTTATGGGAAAGTTGGAGC 1505

Db	1321	GAACCTATTCAGCATGAATAGAAATCTTGGACCTTTCCTTAATGGGAACTTGGAAAC	1380
Qy	1506	ACACTGTTGATTATATTTAAGAAATCAAGACAGACACATATACATCTACCTCACGGA	1565
Db	1381	ACACTGTTGATTATATTTAAGAAATCAAGACAGACACATATACATCTACCTCAGCGA	1440
Qy	1586	ATCACTATGTCCGTCTCTTGTATTCAGAGATTACCAAAAGGTGTAAACATTGAG	1635
Db	1441	ATCACTATGTCCGTCTCTTGTATTCAGAGATTACCAAAAGGTGTAAACATTGAG	1500
Qy	1626	GATTTCCAAATCTGCCAGGAAATATTCAAATATAATGACAGTACGTGTGAACAT	1685
Db	1501	GATTTCCAAATCTGCCAGGAAATATTCAAATATAATGACAGTACGTGTGAACAT	1560
Qy	1686	GGGCCAACTAAATCAGATCCTCGTCCGTGACCCGCTATACCTAGTTGGTTAAATG	1745
Db	1561	GGGCCAACTAAATCAGATCCTCGTCCGTGACCCGCTATACCTAGTTGGTTAAATG	1620
Qy	1746	GAGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCCATCTGCTACAAAGATCTGA	1805
Db	1621	GAGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCCATCTGCTACAAAGATCTGA	1680
Qy	1806	GATCAAAAGAGAAACCGAGTAATGTAGAGACAAGGAATGATCCTGTTTCTGATTT	1885
Db	1681	GATCAAAAGAGAAACCGAGTAATGTAGAGACAAGGAATGATCCTGTTTCTGATTT	1740
Qy	1866	GATGAGAACCGAAGCTGTACCTCACAGAAATATACAGCTTCTCCCAATCAGCT	1925
Db	1741	GATGAGAACCGAAGCTGTACCTCACAGAAATATACAGCTTCTCCCAATCAGCT	1800
Qy	1926	GGAGTCAGCTTGAGATCCAGAGATCCAAAGCTCCAAATCATGACAGCATCAATGC	1985
Db	1801	GGAGTCAGCTTGAGATCCAGAGATCCAAAGCTCCAAATCATGACAGCATCAATGC	1860
Qy	1986	TATGTTTTGATGTTTGGCATGTGTAGTTGTGTCATAGAGTGGCATACGTACATT	2045
Db	1861	TATGTTTTGATGTTTGGCATGTGTAGTTGTGTCATAGAGTGGCATACGTACATT	1920
Qy	2046	CTAAGCATTTGGAGACAGACATGACTTCCTTCTGTCCTGATATACCTTCAA	2105
Db	1921	CTAAGCATTTGGAGACAGACATGACTTCCTTCTGTCCTGATATACCTTCAA	1980
Qy	2106	CACAAAATGCTCTATGGAAGACACACTACCCATATTCATCTCAGAGAAATCTGCTTC	2165
Db	1981	CACAAAATGCTCTATGGAAGACACACTACCCATATTCATCTCAGAGAAATCTGCTTC	2040
Qy	2166	ATGTGATGAAAAACCAAGTCTATGAGATCTTGGGTGCCACAACCTCAGACTTTGGAAC	2235
Db	2041	ATGTGATGAAAAACCAAGTCTATGAGATCTTGGGTGCCACAACCTCAGACTTTGGAAC	2100
Qy	2226	AGAGCATGACCCGCTACTGGAAGGTTTTCAGTTGTGACAACAAACAGTGTATTATAC	2285
Db	2101	AGAGCATGACCCGCTACTGGAAGGTTTTCAGTTGTGACAACAAACAGTGTATTATAC	2160
Qy	2286	GAGGACAGTTATGAAGATATTTACAGATCTTGGCTAGTAAACAAATGCCATTGAACA	2345
Db	2161	GAGGACAGTTATGAAGATATTTACAGATCTTGGCTAGTAAACAAATGCCATTGAACA	2220
Qy	2346	AGAAAGCTTTCGCCAAGAACCAACAGTCTTGAAACGCCATCAAGGGAAATACTGTACT	2405
Db	2221	AGAAAGCTTTCGCCAAGAACCAACAGTCTTGAAACGCCATCAAGGGAAATACTGTACT	2238
Qy	2406	ACCTTGAGTCAGATCAAGAGAAATTTGACTATGATGATACCATATCAGTTGAATGAG	2465
Db	2239	ACCTTGAGTCAGATCAAGAGAAATTTGACTATGATGATACCATATCAGTTGAATGAG	2298
Qy	2466	AAGGAAATTTTGACATTTATGATGAGAGATGAAATACAGAGCCCCGACGCTTCAAAAG	2525
Db	2239	AAGGAAATTTTGACATTTATGATGAGAGATGAAATACAGAGCCCCGACGCTTCAAAAG	2358
Qy	2526	AAAACACGACACTATTTTATTTGCTGACGTGAGAGGCTCTGGGATTTATGGAATGATAC	2585

Dd	2359	AAACACGACACCTATTTTATTGCTGCAGATGAGAGGCTTGGGAATTATGGATGATGATGC	2418
Qy	2586	TCGCCACATGTTCTTAAGAAACAGGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAGAAAGTT	2645
Dd	2419	TCCCCACATGTTCTTAAGAAACAGGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAGAAAGTT	2478
Qy	2646	GTTTTCACGGAATTTACTGATGGCTCTTTTACTCAGCCCTTATACCCTGGAGACTAAAT	2705
Dd	2479	GTTTTCACGGAATTTACTGATGGCTCTTTTACTCAGCCCTTATACCCTGGAGACTAAAT	2538
Qy	2706	GAACATTTGGGACCTCGGGGGCCATATATPAAGGCGAAGTTGGAAGATATATTCATGGTA	2765
Dd	2539	GAACATTTGGGACCTCGGGGGCCATATATPAAGGCGAAGTTGGAAGATATATTCATGGTA	2598
Qy	2766	ACTTTCAGAAATCAGGCGCTCGTCCCTATATCCCTATATTCAGCCTTATTTCTTATGAG	2835
Dd	2599	ACTTTCAGAAATCAGGCGCTCGTCCCTATATCCCTATATTCAGCCTTATTTCTTATGAG	2658
Qy	2836	GAGATCAGAGGCGAAGGACGAACTAGAAAACTTTGTCAAGCTATAGAACCAA	2885
Dd	2659	GAGATCAGAGGCGAAGGACGAACTAGAAAACTTTGTCAAGCTATAGAACCAA	2718
Qy	2886	ACTTACTTTGGAAGGCAACATCATATATGAGCCACCAATGAAGATGAGTTTACATCGCAA	2945
Dd	2719	ACTTACTTTGGAAGGCAACATCATATGAGCCACCAATGAAGATGAGTTTACATCGCAA	2778
Qy	2946	GCTCGGGCTTATTTCTCTGATGTTGACCTGGAAAAAGATGTCACTCAGGCGCTGATTGA	3005
Dd	2779	GCTCGGGCTTATTTCTCTGATGTTGACCTGGAAAAAGATGTCACTCAGGCGCTGATTGA	2838
Qy	3006	CCCCCTTGCTGCGCACACTAAACACACTGAACCTGCTCATATGGAGACAACTGACAGTA	3065
Dd	2839	CCCCCTTGCTGCGCACACTAAACACACTGAACCTGCTCATATGGAGACAACTGACAGTA	2898
Qy	3066	CAGGAATTTGCTGCTGTTT---TCACCATTTGATGAGAACCAAAAGCGATCTTCACT	3122
Dd	2899	CAGGAATTTGCTGCTGTTT---TCACCATTTGATGAGAACCAAAAGCGATCTTCACT	2956
Qy	3123	GAAATATGGAAGAAACCTCAGGGGCTCCCTCGCAATATCAGATGGAAGATCCCACTTT	3182
Dd	2959	GAAATATGGAAGAAACCTCAGGGGCTCCCTCGCAATATCAGATGGAAGATCCCACTTT	3018
Qy	3183	AAAGAAATTTATGCTTCCATGCAATCATATGCTACATATGGAATACACTACCTGGCTTA	3242
Dd	3019	AAAGAAATTTATGCTTCCATGCAATCATATGCTACATATGGAATACACTACCTGGCTTA	3078
Qy	3243	GTAATGGCTCAGATCAAAGGATTCGATGTATCTCTCAGCATGGGACGCAATATAAAC	3302
Dd	3079	GTAATGGCTCAGATCAAAGGATTCGATGTATCTCTCAGCATGGGACGCAATATAAAC	3138
Qy	3303	ATCCATTTCTATCTATTCATGAGGCAATGTTTACGTACGAAAAAAGAGAGATATAA	3366
Dd	3139	ATCCATTTCTATCTATTCATGAGGCAATGTTTACGTACGAAAAAAGAGAGATATAA	3198
Qy	3363	ATGGCAGTGCATCTCTATCCAGGTGTTTTGAGACAGTGGAAATGTTTACATCCCAA	3422
Dd	3199	ATGGCAGTGCATCTCTATCCAGGTGTTTTGAGACAGTGGAAATGTTTACATCCCAA	3258
Qy	3423	GCTGGAATTTGGCGGTGGGAATGCTTATTTGGGAGCATCTACATGCTGGATGAGACA	3482
Dd	3259	GCTGGAATTTGGCGGTGGGAATGCTTATTTGGGAGCATCTACATGCTGGATGAGACA	3318
Qy	3483	CTTTTTCGGGTACACCAATTAAGTGTGACAGATCCCCCTGGGAATGGCTTGGACACATT	3542
Dd	3319	CTTTTTCGGGTACACCAATTAAGTGTGACAGATCCCCCTGGGAATGGCTTGGACACATT	3378
Qy	3543	AGAGATTTTCAGATTACAGCTTCAGGCAATATGAGACATGGGCCCCAAAGCTGGCCAGA	3602
Dd	3379	AGAGATTTTCAGATTACAGCTTCAGGCAATATGAGACATGGGCCCCAAAGCTGGCCAGA	3438
Qy	3603	CTTCATTTATTCGGATCAATCAATGCTGTGGAGCCAAAGAGCCCTTTCTTGGATCAAG	3662
Dd	3439	CTTCATTTATTCGGATCAATCAATGCTGTGGAGCCAAAGAGCCCTTTCTTGGATCAAG	3488

QY 3663 GTGATCTGTGGCCACCAATGATTTATTCACGGCATCAAGCCAGGGTGCCCTCAGAAAG 3722
DB 3499 GTGGATCTGTGGCCACCAATGATTTATTCACGGCATCAAGCCAGGGTGCCCTCAGAAAG 3558
QY 3723 TTCTCAGCCTCAGATCTCTCAGTTTATCATGATGATGTTGATGGGAAGAGTGG 3782
DB 3559 TTCTCAGCCTCAGATCTCTCAGTTTATCATGATGATGTTGATGGGAAGAGTGG 3618
QY 3783 CAGACTTATCGAAGAAATTCACCTGGAACCTTATGTTCTTTTGGCAATGATGATCA 3842
DB 3612 CAGACTTATCGAAGAAATTCACCTGGAACCTTATGTTCTTTTGGCAATGATGATCA 3678
QY 3843 TCTGGGATTAACACAAATATTTTAAACCTTCACATTTATGCTGATACATCGCTTTCAC 3902
DB 3679 TCTGGGATTAACACAAATATTTTAAACCTTCACATTTATGCTGATACATCGCTTTCAC 3738
QY 3903 CCAACTCATTAATAGATTCGACACCTTCGATGAGTGTGATGGGCTGTGATTAAT 3962
DB 3739 CCAACTCATTAATAGATTCGACACCTTCGATGAGTGTGATGGGCTGTGATTAAT 3798
QY 3963 AGTTGACAGATGCGATTTGGGAATGAGAGTAAGCAATATCAGATGACAGATTAAGTCT 4022
DB 3799 AGTTGACAGATGCGATTTGGGAATGAGAGTAAGCAATATCAGATGACAGATTAAGTCT 3858
QY 4023 TCATCTACTTTACCAATATGTTTGGCAGCTGCTCTCTTCAAAAGCTGACTTCACTTC 4082
DB 3859 TCATCTACTTTACCAATATGTTTGGCAGCTGCTCTCTTCAAAAGCTGACTTCACTTC 3918
QY 4083 CAAGGAGAGATGATGCTGAGAGCTCAGGTGATATTCACAAAGAGTGGCTGCAAGTG 4142
DB 3919 CAAGGAGAGATGATGCTGAGAGCTCAGGTGATATTCACAAAGAGTGGCTGCAAGTG 3978
QY 4143 GACTTCCAAAGACATGAAAGTACAGAGATTAAGTCAAGGAGTAAATCTGCTT 4202
DB 3979 GACTTCCAAAGACATGAAAGTACAGAGATTAAGTCAAGGAGTAAATCTGCTT 4038
QY 4203 ACCAGCATGATGTAAGAGTTCCTCATCTCCAGAGTCAAGATGCCATGATGAGT 4262
DB 4039 ACCAGCATGATGTAAGAGTTCCTCATCTCCAGAGTCAAGATGCCATGATGAGT 4098
QY 4263 CTCTTTTTCAGAAATGCAAGAGTAAAGTCTTTCAGGGAATTCAGACTCTTTCACACT 4322
DB 4099 CTCTTTTTCAGAAATGCAAGAGTAAAGTCTTTCAGGGAATTCAGACTCTTTCACACT 4158
QY 4323 GTGGGAAATCTCTAGACCCACCGTTACTGACTCGCTACCTTGAATTCAACCCAGAGT 4382
DB 4159 GTGGGAAATCTCTAGACCCACCGTTACTGACTCGCTACCTTGAATTCAACCCAGAGT 4218
QY 4383 TGGGTGACACCAATGCTGAGAGTGTGAGGAGTGGGAGGACAGAGACCTCTAC 4442
DB 4219 TGGGTGACACCAATGCTGAGAGTGTGAGGAGTGGGAGGACAGAGACCTCTAC 4278

RESULT 3
LOCUS 108643 4281 bp
DEFINITION Sequence 3 from Patent WO 8800831. PAT 14-NOV-1994
ACCESSION 108643
NID 9388649
VERSION 108643.1 GI:588649
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4281)
AUTHORS Pasek, M.P.
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE
POLYPEPTIDES IN HIGH YIELDS
JOURNAL Patent: WO 8800831-A 3 11-FEB-1988;
FEATURES Location/Qualifiers
Source 1..4281

BASE COUNT 1245 a 943 c 946 g 1147 t
ORIGIN
Query Match 86.7%; Score 4190.8; DB 5; Length 4281;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 4270; Conservative 0; Mismatches 2; Indels 48; Gaps 3;
QY 129 GGCACAGAAAGTACTTACTGCTGACAGTGAAGACGTGATGAGGATATATGCAAGTAT 188
DB 4 GGCACAGAAAGTACTTACTGCTGACAGTGAAGACGTGATGAGGATATATGCAAGTAT 63
QY 189 CTCGGTGAAGCTGCTGAGGACCAAGATTTCTCTAGAGTCCAAAATCTTTCCATTC 248
DB 64 CTCGGTGAAGCTGCTGAGGACCAAGATTTCTCTAGAGTCCAAAATCTTTCCATTC 123
QY 249 AACACCTCAGTGTGTACAAAAGACTGTGTTGTAAGATTCAGGATCACCTTTTCAAC 308
DB 124 AACACCTCAGTGTGTACAAAAGACTGTGTTGTAAGATTCAGGATCACCTTTTCAAC 183
QY 309 ATCGTGAAGCCCAAGGCCACCTGATGGTGTGCTAGAGTCCATCCATCCAGCTGAGTT 368
DB 184 ATCGTGAAGCCCAAGGCCACCTGATGGTGTGCTAGAGTCCATCCATCCAGCTGAGTT 243
QY 369 TATGATACAGTGTGATTAACATTAAGATGAGTCCATCCATCCATCCATCCATCCAT 428
DB 244 TATGATACAGTGTGATTAACATTAAGATGAGTCCATCCATCCATCCATCCATCCAT 303
QY 429 GTTGTGTATCTTACTGGAAGACTTCTGAGGAGCTGAATATGATGATCAGCAAGTCAA 488
DB 304 GTTGTGTATCTTACTGGAAGACTTCTGAGGAGCTGAATATGATGATCAGCAAGTCAA 363
QY 489 AGGAGAAAGAAAGATGATAAAGTCTTCCGTGGAGACCAATATGTTGGGAGGCTC 548
DB 364 AGGAGAAAGAAAGATGATAAAGTCTTCCGTGGAGACCAATATGTTGGGAGGCTC 423
QY 549 CTGAAGAAGATGTCATAGGCTCTGACCACTGTCCTTACTACTCATATCTTCT 608
DB 424 CTGAAGAAGATGTCATAGGCTCTGACCACTGTCCTTACTACTCATATCTTCT 483
QY 609 CATGAGACCTGCTGTAAGACTTGAATTCAGGCTCATGAGGCTTACTGATGATGATGA 668
DB 484 CATGAGACCTGCTGTAAGACTTGAATTCAGGCTCATGAGGCTTACTGATGATGATGA 543
QY 669 GAAGGAGTCTGGCCAAAGAAAGACAGACCTTGCACAAATTTATCTACTTTTCT 728
DB 544 GAAGGAGTCTGGCCAAAGAAAGACAGACCTTGCACAAATTTATCTACTTTTCT 603
QY 729 GTATTTGATGAAGGAAAGTGGCACTCAGAAACAAAGAACTCTTGATGACAGATAGG 788
DB 604 GTATTTGATGAAGGAAAGTGGCACTCAGAAACAAAGAACTCTTGATGACAGATAGG 663
QY 789 GATGCTGATCTGCTGGGCTGCTGCTAAATTCACACAGTCAATGTTATGTAACAGG 848
DB 664 GATGCTGATCTGCTGGGCTGCTGCTAAATTCACACAGTCAATGTTATGTAACAGG 723
QY 849 TCTGTG---CCAGGTCTGATGATGATGACACAGAAATCAGTATTTGATGATGATGA 905
DB 724 TCTGTGACACAGGTCTGATGATGATGACACAGAAATCAGTATTTGATGATGATGA 783
QY 906 ATGGGACACACTCTGAAATGCACTCAATATCTCTGAAAGTCAACATTTCTGTGAGG 965
DB 784 ATGGGACACACTCTGAAATGCACTCAATATCTCTGAAAGTCAACATTTCTGTGAGG 843
QY 966 AACATGCGCAGGGTCTGGAATCTCGCCATTAATCTTCTTACTGCTCAACACTC 1025
DB 844 AACATGCGCAGGGTCTGGAATCTCGCCATTAATCTTCTTACTGCTCAACACTC 903
QY 1026 TTGATGACCTTGGACAGTTTCTACTGTTTGTGATATCTCTTCCCAACATGATGAGC 1085
DB 904 TTGATGACCTTGGACAGTTTCTACTGTTTGTGATATCTCTTCCCAACATGATGAGC 963


```

Db 3082 GTATGCTCAGATCAAGGATTGCATGATCTGCTCAGCATGGGACAGCATGAAAC 3141
QY 3303 ATCCATTCTTATTTCACTTTCAGTGGACATGCTTCACTACGAAAAAAGAGATATAAA 3362
Db 3142 ATCCATTCTTATTTCACTTTCAGTGGACATGCTTCACTACGAAAAAAGAGATATAAA 3201
QY 3363 ATGCACTGACATCTCTATCCAGGTGTTTTAGACACTGGAAATGTTACATCCAAA 3422
Db 3202 ATGGCACTGACATCTCTATCCAGGTGTTTTAGACAGTGGAAATGTTACATCCAAA 3261
QY 3423 GCTGGAATTTGGGGGGTGAATGCTTATTTGGGAGCATCTACATGCTGGGATGAGACA 3482
Db 3262 GCTGGAATTTGGGGGGTGAATGCTTATTTGGGAGCATCTACATGCTGGGATGAGACA 3321
QY 3483 CTTTTCGTGTGACAGCAATTAAGTGTCAACCTCCCTGGGAATGCTTTGACACATT 3542
Db 3322 CTTTTCGTGTGACAGCAATTAAGTGTCAACCTCCCTGGGAATGCTTTGACACATT 3381
QY 3543 AGAGATTTTACAGTTACAGTCTCAGGACCAATGAGACAGTGGGCCCCAAAAGTGGCCAGA 3602
Db 3382 AGAGATTTTACAGTTACAGTCTCAGGACCAATGAGACAGTGGGCCCCAAAAGTGGCCAGA 3441
QY 3603 CTTTCATTATCCGGAATCAATCAATGCTGAGACCAAGAGGCCCTTTCTTGGATCAAG 3662
Db 3442 CTTTCATTATCCGGAATCAATCAATGCTGAGACCAAGAGGCCCTTTCTTGGATCAAG 3501
QY 3663 GTGGATCTGTGGACCAATGATTTTACAGGATCAAGGCCAGGGTGGCCCTCAGAG 3722
Db 3502 GTGGATCTGTGGACCAATGATTTTACAGGATCAAGGCCAGGGTGGCCCTCAGAG 3561
QY 3723 TTCTCCAGCCTTACATCTCTCAGTTTATCATGATGATGATGTTGAGGAAGATGG 3782
Db 3562 TTCTCCAGCCTTACATCTCTCAGTTTATCATGATGATGATGTTGAGGAAGATGG 3621
QY 3783 CAGACTTATCGAGAAATTCCTGAGAACCTTATGCTTTTGGCAATGTGATTC 3842
Db 3622 CAGACTTATCGAGAAATTCCTGAGAACCTTATGCTTTTGGCAATGTGATTC 3681
QY 3843 TCTGGATATAAACAATATTTTAAACCTCCCAATTTCTGATACATCCGTTTGCAC 3902
Db 3682 TCTGGATATAAACAATATTTTAAACCTCCCAATTTCTGATACATCCGTTTGCAC 3741
QY 3903 CCAACTATTTATAGCATTCGACGACACTCTTCGATGGAGTTGATGGGCTGTGATTTAAAT 3962
Db 3742 CCAACTATTTATAGCATTCGACGACACTCTTCGATGGAGTTGATGGGCTGTGATTTAAAT 3801
QY 3963 AGTTGACAGCATGCCATTGGGATGAGAGTAAAGCAATATCAGATGCACAGATTACTGCT 4022
Db 3802 AGTTGACAGCATGCCATTGGGATGAGAGTAAAGCAATATCAGATGCACAGATTACTGCT 3861
QY 4023 TCAATCTACTTTACCAATATGTTTGCACCTGGTCTCCCTTCAAAAGCTGACTTCACTTC 4082
Db 3862 TCAATCTACTTTACCAATATGTTTGCACCTGGTCTCCCTTCAAAAGCTGACTTCACTTC 3921
QY 4083 CAAGGAGAGATATGCTTGGAGACCTCAGGTGAATATATCCAAAAGTGGCTGCAAGTG 4142
Db 3922 CAAGGAGAGATATGCTTGGAGACCTCAGGTGAATATATCCAAAAGTGGCTGCAAGTG 3981
QY 4143 GACTTCCAAAGCAATGAAAGTCAAGAGTAACTACCTACGAGTAAATCTCTGCTT 4202
Db 3982 GACTTCCAAAGCAATGAAAGTCAAGAGTAACTACCTACGAGTAAATCTCTGCTT 4041
QY 4203 ACCAGCATGATGTAAGAGTCTCATCTCAGAGTCGAAGATGAGCATGCTGAGACT 4262
Db 4042 ACCAGCATGATGTAAGAGTCTCATCTCAGAGTCGAAGATGAGCATGCTGAGACT 4101
QY 4263 CTCTTTTTCAGAAATGCAAAAGTATTTTCAAGGAAATCAAGACTCTTCAACCT 4322
Db 4102 CTCTTTTTCAGAAATGCAAAAGTATTTTCAAGGAAATCAAGACTCTTCAACCT 4161
QY 4323 GTGGTGAATCTCTAGACCAACGCTTACTGCTGCTACCTTGAATTCACCCCAAGT 4382
Db 4162 GTGGTGAATCTCTAGACCAACGCTTACTGCTGCTACCTTGAATTCACCCCAAGT 4221

```

```

QY 4383 TGGGTCCACCAGATTGCCCTGAGAGATGAGGTTCTGGGCTGGAGGACAGGACCTCTAC 4442
Db 4222 TGGGTCCACCAGATTGCCCTGAGAGATGAGGTTCTGGGCTGGAGGACAGGACCTCTAC 4281

RESULT 4
LOCUS 108641 4551 bp
DEFINITION Sequence 1 from Patent WO 8800831. PAT 14-NOV-1994
ACCESSION 108641
NID 9588647
VERSION 108641.1 GI:588647
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4551)
AUTHORS Pask, M.P.
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR
POLYPEPTIDES IN HIGH YIELDS
JOURNAL Viii:C-Like POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE
FEATURES Patent: WO 8800831-A 1 11-FEB-1988;
SOURCE Location/Qualifiers
BASE COUNT 1345 a 1003 c 1004 g 1198 t 1 others
ORIGIN

Query Match
Best Local Similarity 94.8%; Score 4045.8; DB 5; Length 4551;
Matches 4311; Conservative 0; Mismatches 3; Indels 234; Gaps 3;

QY 129 GCCACCAAGAAATGATCTACTGCTGGTGCAGTGAAGTGTCTGATGGAGTATATGCAAGTAT 188
Db 4 GCCACCAAGAAATGATCTACTGCTGGTGCAGTGAAGTGTCTGATGGAGTATATGCAAGTAT 63
QY 189 CTGGGTGAGCTGCTGTGGAGCCAAAGATTTCTCTAGAGTGCCAAAATCTTTTCCATTTC 248
Db 64 CTGGGTGAGCTGCTGTGGAGCCAAAGATTTCTCTAGAGTGTCTGATGGAGTATATGCAAGTAT 123
QY 249 AACACTGAGTGTGTACAAAAGACTGTTGTGATGATTTACGAGATCCTTTTCAAC 308
Db 124 AACACTGAGTGTGTACAAAAGACTGTTGTGATGATTTACGAGATCCTTTTCAAC 183
QY 309 ATCGCTAAGCCAAAGCCACCTGATGGTGTCTAGAGTCTTACCATCAGGCTGAGTT 368
Db 184 ATCGCTAAGCCAAAGCCACCTGATGGTGTCTAGAGTCTTACCATCAGGCTGAGTT 243
QY 369 TATGATACAGTGTGATTTACATTAAGAAATGCTTCCATCCCTGATCAGGCTTATGCT 428
Db 244 TATGATACAGTGTGATTTACATTAAGAAATGCTTCCATCCCTGATCAGGCTTATGCT 303
QY 429 GTTGGTATCTTACTGGAAGACTCTGAGGAGTGAATATGATGATGATGATGATGATGAT 488
Db 304 GTTGGTATCTTACTGGAAGACTCTGAGGAGTGAATATGATGATGATGATGATGATGAT 363
QY 489 AAGGAGAAAAGATGATAAAGTCTTCCCTGATGGAAGCCATACATATGCTGAGGCTC 548
Db 364 AAGGAGAAAAGATGATAAAGTCTTCCCTGATGGAAGCCATACATATGCTGAGGCTC 423
QY 549 CTGAAGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 608
Db 424 CTGAAGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 483
QY 609 CATGTGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 668
Db 484 CATGTGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 543
QY 669 GAAGGAGTCTGGCCCAAGAAAAGACAGACATTTGCAAAATTTATATCTACTTTTCT 728
Db 544 GAAGGAGTCTGGCCCAAGAAAAGACAGACATTTGCAAAATTTATATCTACTTTTCT 603

```


Query Match	83.7%; Score 4045.8; DB 5; Length 4548;
Best Local Similarity	94.8%; Pred. No. 0;
Db 3844	TGACGTCTCAGATTATATCATATGATATAGCTTTGATGGGAAGAAGTGGCAGACTTATCGCA 3903
Qy 3795	GGAAATTCACATGGAAACCTTAATGCTCTTTTGGCAATGGATTCATCTGGGATAAAA 3854
Db 3904	GGAAATTCACATGGAAACCTTAATGCTCTTTTGGCAATGGATTCATCTGGGATAAAA 3963
Qy 3855	CACAAATATTTTAAACCCCTCAATTTATGCTGATACATCCGTTTGACACCACCTCATTT 3914
Db 3964	CACAAATATTTTAAACCCCTCAATTTATGCTGATACATCCGTTTGACACCACCTCATTT 4023
Qy 3915	AGCAATTCGACGACCTCTTTCGATGAGTGTGATGGGCTGTGATTTAAATAGTTGCACATG 3974
Db 4024	AGCAATTCGACGACCTCTTTCGATGAGTGTGATGGGCTGTGATTTAAATAGTTGCACATG 4083
Qy 3975	CCATTGGGAATGGAAGATTAACCAATATACAGATGCACAGATTACTGCTCATCTTACTTT 4034
Db 4084	CCATTGGGAATGGAAGATTAACCAATATACAGATGCACAGATTACTGCTCATCTTACTTT 4143
Qy 4035	ACCAATATGTTTGGCCACCGCTGCTCCTCCAAAGGTCGACCTCACCTCCAAAGGAGAGT 4094
Db 4144	ACCAATATGTTTGGCCACCGCTGCTCCTCCAAAGGTCGACCTCACCTCCAAAGGAGAGT 4203
Qy 4095	AATGCTTGAGACCTCTCAGGTGATATATCCAAAGAAGTGGCTGCAAGTGGACTTCCAGAAG 4154
Db 4204	AATGCTTGAGACCTCTCAGGTGATATATCCAAAGAAGTGGCTGCAAGTGGACTTCCAGAAG 4263
Qy 4155	ACAATGAAAGTCACAGAGGTAACTACTCTCAGAGAGTAATATCTCTGCTTACACGACATGAT 4214
Db 4264	ACAATGAAAGTCACAGAGGTAACTACTCTCAGAGAGTAATATCTCTGCTTACACGACATGAT 4323
Qy 4215	GTGAAGGAGTTCCTCATCTCCAGCAGCAGTCAGATGGCCATCAGTGGACCTCTTTTTCAG 4274
Db 4324	GTGAAGGAGTTCCTCATCTCCAGCAGCAGTCAGATGGCCATCAGTGGACCTCTTTTTCAG 4383
Qy 4275	AATGGCAAGTAAGAGTTTTTTCAGGGAATCAAGACTCTTTCACACCTGTGTGAACTCT 4334
Db 4384	AATGGCAAGTAAGAGTTTTTTCAGGGAATCAAGACTCTTTCACACCTGTGTGAACTCT 4443
Qy 4335	CTAGACCCACCGTTACTGACTCGTACCTTGAAATTCACCCCAAGAGTTGGGTGCACACG 4394
Db 4444	CTAGACCCACCGTTACTGACTCGTACCTTGAAATTCACCCCAAGAGTTGGGTGCACACG 4503
Qy 4395	ATTGCCCTGAGAGATGAGAGTTCTGGGCTGCGAGGACACAGAGTGGGTGATC 4442
Db 4504	ATTGCCCTGAGAGATGAGAGTTCTGGGCTGCGAGGACACAGAGTGGGTGATC 4551
RESULT 5	
LOCUS 108642	108642 4548 bp PAT 14-NOV-1994
DEFINITION	Sequence 2 from Patent WO 8800831.
ACCSSION	108642
NID	9588648
VERSION	108642.1 GI:588648
KEYWORDS	Unknown.
SOURCE	Unknown.
ORGANISM	Unclassified.
REFERENCE	1 (Bases 1 to 4548)
AUTHORS	Pasek,M.P.
TITLE	DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE POLYPEPTIDES IN HIGH YIELDS
JOURNAL	Patent: WO 8800831-A 2 11-FEB-1988;
FEATURES	Location/Qualifiers
source	1..4548
BASE COUNT	1344 a 1004 c 1002 g 1197 t 1 others
ORIGIN	/organism="unknown"

Db 4321 GTGAAGAGATTCCTCATCTCCAGACAGTCAAGATGGGCATCACTGACACTCTTTTTCAG 4380
QY 4275 AATGCCAAGTAAGTGTTCCTTTCAGGGAATCAAGACTCTTACACCTGTGTGAACTCT 4334
Db 4381 AATGGCAAGTAAGTGTTCCTTTCAGGGAATCAAGACTCTTACACCTGTGTGAACTCT 4440
QY 4335 CTAGACCCACCTGTACTGACTGCTACCTTCCATTCACCCAGAGTGGGTGACACG 4394
Db 4441 CTAGACCCACCTGTACTGACTGCTACCTTCCATTCACCCAGAGTGGGTGACACG 4500
QY 4395 ATTGCCCTGAGAGTGAAGTTCCTGAGGCTGCGAGGACAGACCTCTAC 4442
Db 4501 ATTGCCCTGAGAGTGAAGTTCCTGAGGCTGCGAGGACAGACCTCTAC 4548
RESULT 6
LOCUS AR003710 9009 bp DNA PAT 04-DEC-1998
DEFINITION Sequence 1 from patent US 5744446.
ACCESSION AR003710
VERSION 93864969
KEYWORDS AR003710.1 GI:3964969
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 9009)
AUTHORS Lollar,J.S. and Runge,M.S.
TITLE Hybrid human/animal factor VIII
JOURNAL Patent: US 5744446-A 1 28-APR-1998;
FEATURES
Source 1. 9009 /organism="unknown"
BASE COUNT 2853 a 1907 c 1844 g 2405 t
ORIGIN

Query Match 51.1%; Score 2468.8; DB 5; Length 9009;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2173; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2345 AAGAAGCTTCTCCAGAACCCACAGCTTTGAAGCCATCAAGGGAATTAATCTGTCAC 2404
Db 5106 AAGGCTGTGCTCTCAAAACCCACAGCTTTGAAGCCATCAAGGGAATTAATCTGTCAC 5165
QY 2405 TACTCTTCAGTCAGATCAAGAGGAATGACTATGATGATACCATATCACTTGAATGAA 2464
Db 5166 TACTCTTCAGTCAGATCAAGAGGAATGACTATGATGATACCATATCACTTGAATGAA 5225
QY 2465 GAAGAAGATTTTGACATTTATGATGAGATGAAATCGAGCCCCCGACCTTTCAGAA 2524
Db 5226 GAAGAAGATTTTGACATTTATGATGAGATGAAATCGAGCCCCCGACCTTTCAGAA 5285
QY 2525 GAAGAAGATTTTGACATTTATGATGAGATGAAATCGAGCCCCCGACCTTTCAGAA 2584
Db 5286 GAAGAAGATTTTGACATTTATGATGAGATGAAATCGAGCCCCCGACCTTTCAGAA 5345
QY 2585 CTCCACCATGTTCTAAGAACAGGAGCTCAGAGTGGCAGTCTCCCTCACTTCAAGAAAGT 2644
Db 5346 CTCCACCATGTTCTAAGAACAGGAGCTCAGAGTGGCAGTCTCCCTCACTTCAAGAAAGT 5405
QY 2645 TGTTCCTCAGGAATTTACATGATGAGTCCCTTACTCAGCCCTTATACCGTGGAGAACTAA 2704
Db 5406 TGTTCCTCAGGAATTTACATGATGAGTCCCTTACTCAGCCCTTATACCGTGGAGAACTAA 5465
QY 2705 TGAACATTTGGAGTCTCTGGGCAATATATAGAGAGAGAGTGAAGATTAATCATGTG 2764
Db 5466 TGAACATTTGGAGTCTCTGGGCAATATATAGAGAGAGAGTGAAGATTAATCATGTG 5525
QY 2765 AACTTCAGAAATCAGGCTCTGCTGCTATTCCTTCTATTCCTAGCCTTATTTCTATGA 2824
Db 5526 AACTTCAGAAATCAGGCTCTGCTGCTATTCCTTCTATTCCTAGCCTTATTTCTATGA 5585

QY 2825 GGAAGATCAGAGCGAAGAGGACGAACTAGAAACCTTTGTCAAGCCTAATGAACCAA 2884
Db 5586 GGAAGATCAGAGCGAAGAGGACGAACTAGAAACCTTTGTCAAGCCTAATGAACCAA 5645
QY 2885 AACTTACTTTTGAAGTGAACATCATATGACACCCACTAAAGATGTGATGACGAA 2944
Db 5646 AACTTACTTTTGAAGTGAACATCATATGACACCCACTAAAGATGTGATGACGAA 5705
QY 2945 AGCTGGCCTTATTTCTGTGATGTGACCTGAGAAAGATGTGACCTAGGCTGATTTGG 3004
Db 5706 AGCTGGCCTTATTTCTGTGATGTGACCTGAGAAAGATGTGACCTAGGCTGATTTGG 5765
QY 3005 ACCCTTGTGCTGCTCCACACTAAGACCTGACCTGCTATGAGGAGAACAGTACAGT 3064
Db 5766 ACCCTTGTGCTGCTCCACACTAAGACCTGACCTGCTATGAGGAGAACAGTACAGT 5825
QY 3065 ACAGGAATTTGCTGCTGCTTTCACCATCTTTGATGAGACCAAAAGCTGTACTGACGTA 3124
Db 5826 ACAGGAATTTGCTGCTGCTTTCACCATCTTTGATGAGACCAAAAGCTGTACTGACGTA 5885
QY 3125 AAATATGGAAGAACTCAGAGGCTCCCTGCAATATCCAGATGAGAGATCCACCTTTAA 3184
Db 5886 AAATATGGAAGAACTCAGAGGCTCCCTGCAATATCCAGATGAGAGATCCACCTTTAA 5945
QY 3185 AGAGATTTATGCTCTTCATGCAATGATGCTACATATGATGATACCTAGCTGCTTACT 3244
Db 5946 AGAGATTTATGCTCTTCATGCAATGATGCTACATATGATGATACCTAGCTGCTTACT 6005
QY 3245 AATGCTCAGATCAAGAGATTCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3304
Db 6006 AATGCTCAGATCAAGAGATTCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6065
QY 3305 CCATTCATTTCAATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3364
Db 6066 CCATTCATTTCAATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6125
QY 3365 GGCATCTGACATCTCTATCCAGGCTTTTGAAGACCTGGAAGTACCTTCCAAAGC 3424
Db 6126 GGCATCTGACATCTCTATCCAGGCTTTTGAAGACCTGGAAGTACCTTCCAAAGC 6185
QY 3425 TGAATTTGGCGGTGAGATGCTTATTTGGCGAGATCTACATGCTGGAGTACGACACT 3484
Db 6186 TGAATTTGGCGGTGAGATGCTTATTTGGCGAGATCTACATGCTGGAGTACGACACT 6245
QY 3485 TTTCTGTGATACAGCAATTAAGTGTGACAGTCCCTGGGAATGCTTGTGACACATTAG 3544
Db 6246 TTTCTGTGATACAGCAATTAAGTGTGACAGTCCCTGGGAATGCTTGTGACACATTAG 6305
QY 3545 AGATTTTCAGATTACAGCTTCAGGCAATATGAGAGTGGGCCCCCAAGCTGGCAGACT 3604
Db 6306 AGATTTTCAGATTACAGCTTCAGGCAATATGAGAGTGGGCCCCCAAGCTGGCAGACT 6365
QY 3605 TCATATTCCTGGATCAATCAATGCTGGGAGCAGCAAGGAGCCCTTTGTGATCAAGT 3664
Db 6366 TCATATTCCTGGATCAATCAATGCTGGGAGCAGCAGGAGCCCTTTGTGATCAAGT 6425
QY 3665 GGATCTGTGGCAGCAATGATTAATCAGGCAATCAAGGAGTGGCCCTGACAGAGTT 3724
Db 6426 GGATCTGTGGCAGCAATGATTAATCAGGCAATCAAGGAGTGGCCCTGACAGAGTT 6485
QY 3725 CTCCAGCCTCTACATCTCAGTTCATATCATCATGATGATGATGAGAGAGTGGCA 3784
Db 6486 CTCCAGCCTCTACATCTCAGTTCATATCATCATGATGATGATGAGAGAGTGGCA 6545
QY 3785 GACTATTCAGAGAAATTCACATGAGAACTTAATGATGCTTCTTGGCAATGATGATTC 3844
Db 6546 GACTATTCAGAGAAATTCACATGAGAACTTAATGATGCTTCTTGGCAATGATGATTC 6605
QY 3845 TGGATTAAGCAATATTTTAACTTCACATTAATGCTGATACATCCGTTTGACACC 3904
Db 6606 TGGATTAAGCAATATTTTAACTTCACATTAATGCTGATACATCCGTTTGACACC 6665
QY 3905 AACTCATATAGCAATTCGACAGACTCTTGGCATGAGAGTGGCTGTGATTTAAATAG 3964

Dd	6666	AACTATTATATAGCAATTCGACACACTCTCCGATGAGGTTGATGGCTGTGATTTAAATAG	67255
Oy	3965	TTGCAGCAATGCCATTTGGGAATGGAGATTAAGCAAAATACAGATGCACATTAACGTTC	4024
Dd	6726	TTGCAGCAATGCCATTTGGGAATGGAGATTAAGCAAAATACAGATGCACATTAACGTTC	6785
Oy	4025	ATCTACCTTTACCAATATGTTTGGCACCTGGTCTCCTTCAAAGCTCGACTTCACCTCA	4084
Dd	6786	ATCTACCTTTACCAATATGTTTGGCACCTGGTCTCCTTCAAAGCTCGACTTCACCTCA	6845
Oy	4085	AGGGAGGAGTAATGCTCGAGACCTCAAGTGAAATTAATCCAAAAGAGTGGCTCAAGTGA	4144
Dd	6846	AGGGAGGAGTAATGCTCGAGACCTCAAGTGAAATTAATCCAAAAGAGTGGCTCAAGTGA	6905
Oy	4145	CTTCAGAGAAGCAATGAAAGTCAGAGAGTAATCTACTCAGGGAGTAATAATCTCTCTTAC	4204
Dd	6906	CTTCAGAGAAGCAATGAAAGTCAGAGAGTAATCTACTCAGGGAGTAATAATCTCTCTTAC	6965
Oy	4205	CAGCATGTATGTGAAGAGTTCCTCATCTCCACAGCTCAAGAATGGGCATCAGTGGACTCT	4264
Dd	6966	CAGCATGTATGTGAAGAGTTCCTCATCTCCACAGCTCAAGAATGGGCATCAGTGGACTCT	7025
Oy	4265	CTTTTTCAGATGCGCAAGTAAGGTTTTTTCAGGAAATCAAGACTCCTTCACACTGT	4324
Dd	7026	CTTTTTCAGATGCGCAAGTAAGGTTTTTTCAGGAAATCAAGACTCCTTCACACTGT	7085
Oy	4325	GGTGAACCTCTATAGCCACCGTTACTGACTGCTACCTTCCAAATTCACCCCAAGTGT	4384
Dd	7086	GGTGAACCTCTATAGCCACCGTTACTGACTGCTACCTTCCAAATTCACCCCAAGTGT	7145
Oy	4385	GGTGACACAGATTGCGCTCGAGATGAGAGTTGCGGCTCGAGGACAGAGACCTCTACTG	4444
Dd	7146	GGTGACACAGATTGCGCTCGAGATGAGAGTTGCGGCTCGAGGACAGAGACCTCTACTG	7205
Oy	4445	AGGGTGGCCACTGCGAGCACTGCCACTGCCGCTACTCTCCTCTCAGCTCCAGGGCAG	4504
Dd	7206	AGGGTGGCCACTGCGAGCACTGCCACTGCCGCTACTCTCCTCTCAGCTCCAGGGCAG	7265
Oy	4505	TGTCCCTCCCGGGCTTCCCTTACCTTGTGTGTAATCCTTGACAGACACTGCTTGAG	4564
Dd	7266	TGTCCCTCCCGGGCTTCCCTTACCTTGTGTGTAATCCTTGACAGACACTGCTTGAG	7325
Oy	4565	CTCTCGTAATTAACATATCATCAGTCTCGATTTCTTTGGTGGGGGCCAGAGAGGTGCAT	4624
Dd	7326	CTCTCGTAATTAACATATCATCAGTCTCGATTTCTTTGGTGGGGGCCAGAGAGGTGCAT	7385
Oy	4625	CCAAATTAACCTTAACCTTACCTTAATTTTCTGACAGCTGCTCCAGATTAATCTCTCTTCC	4684
Dd	7386	CCAAATTAACCTTAACCTTACCTTAATTTTCTGACAGCTGCTCCAGATTAATCTCTCTTCC	7445
Oy	4685	AAATTAACCTAGGCAAAAAGAGTGAAGAAACCTGCATGAAAGCATCTCCCTGGAAA	4744
Dd	7446	AAATTAACCTAGGCAAAAAGAGTGAAGAAACCTGCATGAAAGCATCTCCCTGGAAA	7505
Oy	4745	GTTAGGCGCTCTCAGAGTCAACACTTCTCTGTGTGTAAGAAAATGTGATGAATCTTG	4804
Dd	7506	GTTAGGCGCTCTCAGAGTCAACACTTCTCTGTGTGTAAGAAAATGTGATGAATCTTG	7565
Oy	4805	AAAAAGATATTTATGATGTT 4824	
Dd	7566	AAAAAGATATTTATGATGTT 7585	

RESULT	7		PAT	20-DEC-1996
I31901		I31901	9009 bp DNA	
LOCUS		Sequence 3 from patent US 5583209.		
DEFINITION		I31901		
ACCESSION		G1822692		
NID		I31901.1		
VERSION		GI:1822692		
KEYWORDS				
SOURCE		Unknown.		

ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 9009)
TITLE	Lollar,J. S. and Runge,M. S.
JOURNAL	Hybrid human/porcine factor VII
FEATURES	Patent: US 5583209 A 3 10-DEC-1996; Location/Qualifiers
Source	1..9009
BASE COUNT	2853 a
ORIGIN	1907 c 1844 g 2405 t
	/organism="unknown"

Query Match	51.1%	Score 2468.8	DB 5	Length 9009
Best Local Similarity	99.7%	Prod. No. 0		
Matches 2473; Conservative	0	Mismatches 7	Indels 0	Gaps 0

QY	2345	AAGAGCGTTCGCCAGAAACCCACACACTCTTGAACACCCATCAACGGGAAATTAACTCGAC	2405
Db	5106	AAGCGTGTGCTCAGAAAACCCACACACTCTTGAACACCCATCAACGGGAAATTAACTCGAC	5165
QY	2405	TACTCTTCAGTCAGATCAGAGGAAATTGACTATGATGTATACCATATCAGTTGAAATGAA	2465
Db	5166	TACTCTTCAGTCAGATCAGAGGAAATTGACTATGATGTATACCATATCAGTTGAAATGAA	5226
QY	2465	GAAAGAGATTTTGGACATTTATATAGATGAGGAAATTCAGACCCCCCGAGCTTCCAAA	2526
Db	5226	GAAAGAGATTTTGGACATTTATATATAGATGAGGAAATTCAGACCCCCCGAGCTTCCAAA	5286
QY	2525	GA AAAACAGCAGACTATTTTATTTGCTCAGTGGAGAGGCTCTGGATTTATGAGGATGATAG	2586
Db	5286	GA AAAACAGCAGACTATTTTATTTGCTCAGTGGAGAGGCTCTGGATTTATGAGGATGATAG	5346
QY	2585	CTCCCCAGACTGTTCTTAAGAAACAGGGCTCAGAGTGGCAGCTGTCCCTCAGTTCAAGAAAGT	2644
Db	5346	CTCCCCAGACTGTTCTTAAGAAACAGGGCTCAGAGTGGCAGCTGTCCCTCAGTTCAAGAAAGT	5405
QY	2645	TGTTTTCCAGGAATTTACTGATAGGCTGCTTTTACTAGCCCTTATACCGTGGAGACTTAA	2704
Db	5406	TGTTTTCCAGGAATTTACTGATAGGCTGCTTTTACTAGCCCTTATACCGTGGAGACTTAA	5465
QY	2705	TGAACATTTGGGACTCCTTGGGGCCATATATTAAGAGCAGAGTTGGAAGATTAATATCATGT	2764
Db	5466	TGAACATTTGGGACTCCTTGGGGCCATATATTAAGAGCAGAGTTGGAAGATTAATATCATGT	5525
QY	2765	AACCTTCAGAAATCAGAGGCTCTGCTGCTCCATATCTTACTTATTTACCTTATTTCTTATGA	2824
Db	5526	AACCTTCAGAAATCAGAGGCTCTGCTGCTCCATATCTTACTTATTTACCTTATTTCTTATGA	5585
QY	2825	GGAAGATTCAGAGGCAAGAGAGCAGAAACCTTGAAAAAACTTTGTCAGACCTTAATGAAACCA	2884
Db	5586	GGAAGATTCAGAGGCAAGAGAGCAGAAACCTTGAAAAAACTTTGTCAGACCTTAATGAAACCA	5645
QY	2885	AACCTTACTTTTGGAAAGTCCAAACATCATATGTGCACCCACTAAAGATGAGTTGACTGCA	2944
Db	5646	AACCTTACTTTTGGAAAGTCCAAACATCATATGTGCACCCACTAAAGATGAGTTGACTGCA	5705
QY	2945	AGCCTGGGCTTATTTCTGATAGTTGACCTGGAAAAAATGTGCACCTCAGGCTGATATGG	3004
Db	5706	AGCCTGGGCTTATTTCTGATAGTTGACCTGGAAAAAATGTGCACCTCAGGCTGATATGG	5765
QY	3005	ACCCCTTCTGTGTCGCCACACTACACACTGAAACCTCTCATATGGAGACAAAGTACAGT	3064
Db	5766	ACCCCTTCTGTGTCGCCACACTACACACTGAAACCTCTCATATGGAGACAAAGTACAGT	5825
QY	3065	ACAGGAATTTGCTCTGTGTTTTTACCATCTTTGATGAGACCAAAGCTGTACTTCACTGA	3124
Db	5826	ACAGGAATTTGCTCTGTGTTTTTACCATCTTTGATGAGACCAAAGCTGTACTTCACTGA	5885
QY	3125	AAATATGAAAGAAACCTGAGGGGCTCCCTGCAATTTCCAGATGAGAAATCCCATTTTAA	3184
Db	5886	AAATATGAAAGAAACCTGAGGGGCTCCCTGCAATTTCCAGATGAGAAATCCCATTTTAA	5945

3185 AGAATATATCGTCCATCATCATATGCTATGATGATACCTACTGCTTACT 3244
|||||
5946 AGAGAAATATCGCTTCATCATCATATGCTATGATGATACCTACTGCTTACT 6005
|||||
3245 AATGGCTCAGAGATCAAGAGATGATGATGATGATGATGATGATGATGATGAT 3304
|||||
6006 AATGGCTCAGAGATCAAGAGATGATGATGATGATGATGATGATGATGATGAT 6065
|||||
3305 CCATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 3364
|||||
6066 CCATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 6125
|||||
3365 GGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3424
|||||
6126 GGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6185
|||||
3425 TGAATTTGCGGGGTGGAATGCTTATTTGCGAGCATCTACATGCTGGATGAGCAGCT 3484
|||||
6186 TGAATTTGCGGGGTGGAATGCTTATTTGCGAGCATCTACATGCTGGATGAGCAGCT 6245
|||||
3485 TTTTCTGGGTATCAGCAATATGATGATGATGATGATGATGATGATGATGATGAT 3544
|||||
6246 TTTTCTGGGTATCAGCAATATGATGATGATGATGATGATGATGATGATGATGAT 6305
|||||
3545 AGATTTTACATTTACAGCTTACAGCAATATGATGATGATGATGATGATGATGAT 3604
|||||
6306 AGATTTTACATTTACAGCTTACAGCAATATGATGATGATGATGATGATGATGAT 6365
|||||
3605 TCATTTATTCGGATTCATCATCATGCTGAGCACCAGAGAGAGAGAGAGAGAGAG 3664
|||||
6366 TCATTTATTCGGATTCATCATCATGCTGAGCACCAGAGAGAGAGAGAGAGAGAG 6425
|||||
3665 GGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3724
|||||
6426 GGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6485
|||||
3725 CTCAGAGCTTACATCTCTAGATTTATCATCATGATGATGATGATGATGATGATGAT 3784
|||||
6486 CTCAGAGCTTACATCTCTAGATTTATCATCATGATGATGATGATGATGATGATGAT 6545
|||||
3785 GACTTATTCAGAGAAATTCAGTGAACCTTAAATGCTTCTTTGCAATGATGATGATC 3844
|||||
6546 GACTTATTCAGAGAAATTCAGTGAACCTTAAATGCTTCTTTGCAATGATGATGATC 6605
|||||
3845 TGGGATTAACCAATATTTTAACTTCAATATTTGCTGATGATGATGATGATGATGAT 3904
|||||
6606 TGGGATTAACCAATATTTTAACTTCAATATTTGCTGATGATGATGATGATGATGAT 6665
|||||
3905 AACTCATTAAGCATTCGAGCACTCTTGCATGAGTGTATGATGATGATGATGATGATGAT 3964
|||||
6666 AACTCATTAAGCATTCGAGCACTCTTGCATGAGTGTATGATGATGATGATGATGATGAT 6725
|||||
3965 TTGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4024
|||||
6726 TTGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6785
|||||
4025 ATCTTACTTATTAACCAATATTTGCTGATGATGATGATGATGATGATGATGATGAT 4084
|||||
6786 ATCTTACTTATTAACCAATATTTGCTGATGATGATGATGATGATGATGATGATGAT 6845
|||||
4085 AGGAGAGATTAAGCTGAGCACTCTGAGTGAATATTCAGAGAGAGAGAGAGAGAGAGAG 4144
|||||
6846 AGGAGAGATTAAGCTGAGCACTCTGAGTGAATATTCAGAGAGAGAGAGAGAGAGAGAG 6905
|||||
4145 CTTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4204
|||||
6906 CTTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6965
|||||
4205 CAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4264
|||||
6966 CAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7025
|||||
4265 CTTTCTTCAAGATGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4324
|||||

|||||
Db 7026 CTTTCTTCAAGATGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7085
|||||
Qy 4325 GGTGAACCTCTAGAGCAACCGTTACTGATGATGATGATGATGATGATGATGATGATGAT 4384
|||||
Db 7086 GGTGAACCTCTAGAGCAACCGTTACTGATGATGATGATGATGATGATGATGATGATGAT 7145
|||||
Qy 4385 GGTGAACCTCTAGAGCAACCGTTACTGATGATGATGATGATGATGATGATGATGATGAT 4444
|||||
Db 7146 GGTGAACCTCTAGAGCAACCGTTACTGATGATGATGATGATGATGATGATGATGATGAT 7205
|||||
Qy 4445 AGGAGAGATTAAGCTGAGCACTCTGAGTGAATATTCAGAGAGAGAGAGAGAGAGAGAG 4504
|||||
Db 7206 AGGAGAGATTAAGCTGAGCACTCTGAGTGAATATTCAGAGAGAGAGAGAGAGAGAGAG 7265
|||||
Qy 4505 TGTCCCTCCCTGCTTGCCTTACTGATGATGATGATGATGATGATGATGATGATGATGAT 4564
|||||
Db 7266 TGTCCCTCCCTGCTTGCCTTACTGATGATGATGATGATGATGATGATGATGATGATGAT 7325
|||||
Qy 4565 CCTCCGATTTACTATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4624
|||||
Db 7326 CCTCCGATTTACTATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7385
|||||
Qy 4625 CCAATTTAACTTACTCTTACTATTTTGCAGCTGCTCCAGATTTACTCTTCTTCC 4684
|||||
Db 7386 CCAATTTAACTTACTCTTACTATTTTGCAGCTGCTCCAGATTTACTCTTCTTCC 7445
|||||
Qy 4685 AATATACTAGGCAAAAGAGATGAGAGAACTCATGAGAAATTTCTTCCCTGAAA 4744
|||||
Db 7446 AATATACTAGGCAAAAGAGATGAGAGAACTCATGAGAAATTTCTTCCCTGAAA 7505
|||||
Qy 4745 GTTAGGCTCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4804
|||||
Db 7506 GTTAGGCTCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7565
|||||
Qy 4805 AAAAGATTTATGATGTT 4824
|||||
Db 7566 AAAAGATTTATGATGTT 7585
|||||

RESULT 8
163424
LOCUS 163424 9009 bp DNA PAT 26-SEP-1997
DEFINITION Sequence 1 from patent US 5663060.
ACCESSION 163424
NID 92480997
VERSION 163424.1 GI:2480997 /
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 9009)
AUTHORS Lollar J.S. and Runge M.S.
TITLE Hybrid human/animal factor VIII
JOURNAL Patent: US 5663060-A 1 02-SEP-1997;
FEATURES
source 1. 9009
BASE COUNT 2853 a 1907 c 1844 g 2405 t
ORIGIN
Query Match 51.1%; Score 2468.8; DB 5; Length 9009;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2473; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 2345 AAGAGCTTCTCCAGAGACCAACCAAGTCTGAAACGCCATCAACGGGAATTAAGTGTAC 2404
|||||
Db 5106 AAGGCTGTGCTCAAAACCCAGAGCTTGAAGCCATCAACGGGAATTAAGTGTAC 5165
|||||
Qy 2405 TACTCTTCACTGATCAAGAGAAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 2464
|||||
Db 5166 TACTCTTCACTGATCAAGAGAAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 5225
|||||

QY 2465 GAGGAGATTTTACATTATATGATGAGTGAATTCAGAGCCCGGAGCTTTCAAA 2524
|||||
UB 5226 GAAGGAGATTTTTCATTTATATGATGAGTGAATTCAGAGCCCGGAGCTTTCAAA 5285
QY 2525 GAAAGACACACATTTTATTTGCTGAGTGGAGAGGCTCTGGGATTTATGGATGAGT 2584
|||||
DB 5286 GAAAGACACACATTTTATTTTCTGCTGAGTGGAGAGGCTCTGGGATTTATGGATGAGT 5345
QY 2585 CTCCCAACATGTTCTTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGT 2644
5346 CTCCCAACATGTTCTTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGT 5405
QY 2645 TGTTCCTCAGGATTTACGATGGCTCTTCTACTGAGCCCTTATACCTGGAGAGTAA 2704
5406 TGTTCCTCAGGATTTACTGATGGCTCTTCTACTGAGCCCTTATACCTGGAGAGTAA 5465
DB 2705 TGAACATTTGGGAGCTCTCTGGGCTATATTAAGAGCAGAAATGAAGATATATCATGT 2764
5466 TGAACATTTGGGAGCTCTCTGGGCTATATTAAGAGCAGAAATGAAGATATATCATGT 5525
QY 2765 AACTTTCAAAATTCAGGCTCTCTGCTCTATTTCTTATCTAGCCTTATTTCTATGA 2824
5526 AACTTTCAAAATTCAGGCTCTCTGCTCTATTTCTTATCTAGCCTTATTTCTATGA 5585
QY 2825 GGAAGATCAGAGCAGAGCAGAGAACTTGAAGAAACTTGTCAAGCCCTAATGAACCA 2884
5586 GGAAGATCAGAGCAGAGCAGAGAACTTGAAGAAACTTGTCAAGCCCTAATGAACCA 5645
QY 2885 AACTTACTTTGGAAAGTGAACATCATATGACACCACCTAAGATGAGTTGACTGCA 2944
5646 AACTTACTTTGGAAAGTGAACATCATATGACACCACCTAAGATGAGTTGACTGCA 5705
QY 2945 ACCCTGGGCTTATTTCTGATGTGACCTGGAAAAAGTGTGCACTAGGCTTATTTGG 3004
5706 ACCCTGGGCTTATTTCTGATGTGACCTGGAAAAAGTGTGCACTAGGCTTATTTGG 5765
DB 3005 ACCCTTGTGCTCTGCTCACTACACACTGAACCTGCTCATGGAGACAAATGACAGT 3064
5766 ACCCTTGTGCTCTGCTCACTACACACTGAACCTGCTCATGGAGACAAATGACAGT 5825
QY 3065 ACAGAAATTTGCTCTGTTTTCATCATCTTTGATGAGACCAAAAGCTGTGACTCATGA 3124
5826 ACAGAAATTTGCTCTGTTTTCATCATCTTTGATGAGACCAAAAGCTGTGACTCATGA 5885
DB 3125 AAATATGAAAGAACTGAGGGCTCCCTGCAATTCAGAGATGAAAGATCCCACTTTAA 3184
5886 AAATATGAAAGAACTGAGGGCTCCCTGCAATTCAGAGATGAAAGATCCCACTTTAA 5945
QY 3185 AGAGATTTATGCTTCCATGCAATCATGCTACATTAATGATACACTGCTTATG 3244
5946 AGAGATTTATGCTTCCATGCAATCATGCTACATTAATGATACACTGCTTATG 6005
QY 3245 AATGGCTAGATCAAAAGATTCGATGCTATCTGCTGAGCATGGGACCAATGAACAT 3304
6006 AATGGCTAGATCAAAAGATTCGATGCTATCTGCTGAGCATGGGACCAATGAACAT 6065
DB 3305 CCATTTATTTCAATTCAGTGAATGTGCTGATGAGAAAAAAGAGAGTATTAAT 3364
6066 CCATTTATTTCAATTCAGTGAATGTGCTGATGAGAAAAAAGAGAGTATTAAT 6125
QY 3365 GGCACGTACATCTCTATCCAGGTGTTTTGAAGACAGTGAAGATGTTACCATCAAGC 3424
6126 GGCACGTACATCTCTATCCAGGTGTTTTGAAGACAGTGAAGATGTTACCATCAAGC 6185
DB 3425 TGAATTTGGGGGGGAGATGCTTATTTGGGAGCATCTACATGCTGGAGTGAACACT 3484
6186 TGAATTTGGGGGGGAGATGCTTATTTGGGAGCATCTACATGCTGGAGTGAACACT 6245
QY 3485 TTTTCTGTGTACAGCAATAGTGTGAGACTCCCTGGGAATGGCTTCTGGACATTTAG 3544
6246 TTTTCTGTGTACAGCAATAGTGTGAGACTCCCTGGGAATGGCTTCTGGACATTTAG 6305

QY 3545 AGATTTTCAGATTCAGGCTTCAGGACAAATATGACAGTGGGCCCCAAAGCTGGCCAGCT 3604
|||||
DB 6306 AGATTTTCAGATTCAGGCTTCAGGACAAATATGACAGTGGGCCCCAAAGCTGGCCAGCT 6365
QY 3605 TCATTTATCCGGATCAATCAATGCTGGAGGACCAAGAGGCCCTTTCTTGGATCAAGT 3664
6366 TCATTTATCCGGATCAATCAATGCTGGAGGACCAAGAGGCCCTTTCTTGGATCAAGT 6425
QY 3665 GGATCTGTGGCACAATGATTTATTCAGGGCATCAAGCCAGGGTCCCTCAGAAAGT 3724
6426 GGATCTGTGGCACAATGATTTATTCAGGGCATCAAGCCAGGGTCCCTCAGAAAGT 6485
DB 3725 CTCCAGCCTTACATCTCTCAGTTTATCATGATATATAGTCTTGGAGGAGAAAGTGGCA 3784
6486 CTCCAGCCTTACATCTCTCAGTTTATCATGATATATAGTCTTGGAGGAGAAAGTGGCA 6545
QY 3785 GACTTATGAGGAATTTCCACATGAGTGAACCTTATATGCTCTTTGGCAATGTGATTCATC 3844
6546 GACTTATGAGGAATTTCCACATGAGTGAACCTTATATGCTCTTTGGCAATGTGATTCATC 6605
DB 3845 TGGATTAACACAAATATTTTAACTCCCAATATATGCTGATACATTCGTTTGCACCC 3904
6606 TGGATTAACACAAATATTTTAACTCCCAATATATGCTGATACATTCGTTTGCACCC 6665
QY 3905 AACTCATTTATGCAATTCGACAGACTCTTTCGATGAGAGTGTGATGATTTAAATAG 3964
6666 AACTCATTTATGCAATTCGACAGACTCTTTCGATGAGAGTGTGATGATTTAAATAG 6725
DB 3965 TTGCAGATGCAATTTGGGAATGGAGATTAAGCAATTCAGATGACACAGTACTGCTTC 4024
6726 TTGCAGATGCAATTTGGGAATGGAGATTAAGCAATTCAGATGACACAGTACTGCTTC 6785
QY 4025 ATCTTACTTTACCAATATGTTTGGCACTGCTGCTCTTCAAAAGCTGCAATTCACCTCA 4084
6786 ATCTTACTTTACCAATATGTTTGGCACTGCTGCTCTTCAAAAGCTGCAATTCACCTCA 6845
DB 4085 AGGAGAGATTAATGCTTGGAGAGCTCAGGTGAATATCCAAAGAGTGGCTGCAAGTGA 4144
6846 AGGAGAGATTAATGCTTGGAGAGCTCAGGTGAATATCCAAAGAGTGGCTGCAAGTGA 6905
QY 4145 CTTCAGAAAGCAATGAATGAATCAGAGAGTAACTCAGAGGAGTAAATCTGCTTAC 4204
6906 CTTCAGAAAGCAATGAATGAATCAGAGAGTAACTCAGAGGAGTAAATCTGCTTAC 6965
DB 4205 CAGCATATGTAAGAGAGTCTCTCATCTTCACAGCACTCAAGATGGCCATCAGTGA 4264
6966 CAGCATATGTAAGAGAGTCTCTCATCTTCACAGCACTCAAGATGGCCATCAGTGA 7025
QY 4265 CTTTTCAGAAATGGCAAGTAAAGTTTTTCAGGAAATCAAGACTCCTTACACCTGT 4324
7026 CTTTTCAGAAATGGCAAGTAAAGTTTTTCAGGAAATCAAGACTCCTTACACCTGT 7085
DB 4325 GGTGAACCTCTAGACCAACCGTTACTGACTGCTCTCTTCAATTCACCCCAAGATTG 4384
7086 GGTGAACCTCTAGACCAACCGTTACTGACTGCTCTCTTCAATTCACCCCAAGATTG 7145
QY 4385 GGTGACACAGATGCTGCTGAGATGAGAGTCTGGGCTCTGAGGACAGAGACTTACTG 4444
7146 GGTGACACAGATGCTGCTGAGATGAGAGTCTGGGCTCTGAGGACAGAGACTTACTG 7205
DB 4445 AGGATGGCACTGACAGCTGCTGCACTGCGTCACTCCCTCCCTCAGCTCAAGGCGAG 4504
7206 AGGATGGCACTGACAGCTGCTGCACTGCGTCACTCCCTCCCTCAGCTCAAGGCGAG 7265
QY 4505 TGTCCCTCCCTGGCTGCTTCTTACTTTGTGCTAAATCTTGAAGACACTGCTTGAAG 4564
7266 TGTCCCTCCCTGGCTGCTTCTTACTTTGTGCTAAATCTTGAAGACACTGCTTGAAG 7325
DB 4565 CTTCTGGAATTAATCTATCATGCTCTGATTTCTTGGTGGGGGCGCAGAGGCTGAT 4624
7326 CTTCTGGAATTAATCTATCATGCTCTGATTTCTTGGTGGGGGCGCAGAGGCTGAT 7385
QY 4625 CCAATTTAATTAATCTTACTATTTTCTGACAGTGTGCTCCAGATTTACTCTTCTCC 4684

|||||
Db 7386 CCAATTACTACTACTACTATTTCTGACGCTGCCACAGATTAATCTCTCTCC 7445
OY 4665 AATATTAAGCAAAAGAAAGTGAAGAAACCTGCATGAAGATTTCTCCGAAAA 4744
Db 7446 AATATTAAGCAAAAGAAAGTGAAGAAACCTGCATGAAGATTTCTCCGAAAA 7505
OY 4745 GTTAGGCTCTGAGAGTCCACCTCTCTGTAGAAAATATGTAGAACTTTG 4804
Db 7506 GTTAGGCTCTGAGAGTCCACCTCTCTGTAGAAAATATGTAGAACTTTG 7565
OY 4805 AAAAGATATTTATGATGT 4824
Db 7566 AAAAGATATTTATGATGT 7585

RESULT 9
171409 171409 8967 bp DNA PAT 23-DEC-1997
LOCUS DEFINITION Sequence 1 from patent US 5681746.
ACCESSION 171409
ID 93007544
VERSION 171409.1 GI:3007544
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 8967)
AUTHORS Bodner,M., De Polo,N.J., Chang,S., Hsu,D.,Chl-Tang and Respses,J.G.
TITLE Retrieval of delivery of full length factor VIII
JOURNAL Patent: US 5681746-A 1 28-OCT-1997;
FEATURES
source 1. 8967 Location/Qualifiers
BASE COUNT 2841 a 1898 c 1833 g 2395 t
ORIGIN

Query Match 51.1%; Score 2468.8; DB 5; Length 8967;
Local Similarity 99.7%; Pred. No. 0;
Matches 2473; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 2825 GGAAGATGAGGCGAAGAGGAGAAACCTAGAAAACCTTTGTCAGCCTAATGAAACCA 2884
Db 5545 GGAAGATGAGGCGAAGAGGAGAAACCTAGAAAACCTTTGTCAGCCTAATGAAACCA 5604
OY 2885 AACTTACTTTTGAAGAGCAACATCATATGAGCACCCTAAAGATTTGATGTCGAA 2944
Db 5605 AACTTACTTTTGAAGAGCAACATCATATGAGCACCCTAAAGATTTGATGTCGAA 5664
OY 2945 AGCCTGGCTATTTCTGTATGTTGACCTGGAAAAAGATGTGCATCAGGCTGATTGG 3004
Db 5665 AGCCTGGCTATTTCTGTATGTTGACCTGGAAAAAGATGTGCATCAGGCTGATTGG 5724
OY 3005 ACCCTTGTGCTGCTCCACATCAACACATGAAACCTCTCATGGGAGACAGTACACT 3064
Db 5725 ACCCTTGTGCTGCTCCACATCAACACATGAAACCTCTCATGGGAGACAGTACACT 5784
OY 3065 ACAGAAATTTGCTGTTTTCACCATCTTGATGAGACCAAAACCTGTTACTGACGTA 3124
Db 5785 ACAGAAATTTGCTGTTTTCACCATCTTGATGAGACCAAAACCTGTTACTGACGTA 5844
OY 3125 AAATATGAAAGAAACCTGAGGCTCCCTGCAATATCCAGATGAAAGATCCCACTTTAA 3184
Db 5845 AAATATGAAAGAAACCTGAGGCTCCCTGCAATATCCAGATGAAAGATCCCACTTTAA 5904
OY 3185 AGAGATTAATGCTTCCATGCATATGATGCTACATTAATGATACACTGCTGCTTACT 3244
Db 5905 AGAGATTAATGCTTCCATGCATATGATGCTACATTAATGATACACTGCTGCTTACT 5964
OY 3245 AATGCTCAGATGCAAAAGATTCGATGATGCTGATGCTGACATGGGACCAATGAAACAT 3304
Db 5965 AATGCTCAGATGCAAAAGATTCGATGATGCTGATGCTGACATGGGACCAATGAAACAT 6024
OY 3305 CCATCTATTTCTTTCAGTGCATGTCTCCTGACATGCAAAAAAGAGATATTAAT 3364
Db 6025 CCATCTATTTCTTTCAGTGCATGTCTCCTGACATGCAAAAAAGAGATATTAAT 6084
OY 3365 GGCATGTACATCTCTATCCAGGCTTTTGAACAGATGCAAAATGTTACCTCCAAAGC 3424
Db 6085 GGCATGTACATCTCTATCCAGGCTTTTGAACAGATGCAAAATGTTACCTCCAAAGC 6144
OY 3425 TGGAAATTTGGGCGGTGATGCTTATGGGAGATGTACATGCTGGGATGAGCACT 3484
Db 6145 TGGAAATTTGGGCGGTGATGCTTATGGGAGATGTACATGCTGGGATGAGCACT 6204
OY 3485 TTTTCTGTGTCAGCAATTAAGTGTCAAGCTCCCTGGGAATGGCTTTGACACATTAG 3544
Db 6205 TTTTCTGTGTCAGCAATTAAGTGTCAAGCTCCCTGGGAATGGCTTTGACACATTAG 6264
OY 3545 AGATTTTTCAGATTAAGCTTTCAGCAATATGAGACATGAGGCCCCCAAAAGCTGCCAGACT 3604
Db 6265 AGATTTTTCAGATTAAGCTTTCAGCAATATGAGACATGAGGCCCCCAAAAGCTGCCAGACT 6324
OY 3605 TCATTAATCCGATCAATCAATGCTGAGGAGCAAGAGGCTTTTGTGATCAAGT 3664
Db 6325 TCATTAATCCGATCAATCAATGCTGAGGAGCAAGAGGCTTTTGTGATCAAGT 6384
OY 3665 GGATCTGTTGGCACCATGATTAATTCAGGCAATCAAGGCTGCCCTCAGAAAGT 3724
Db 6385 GGATCTGTTGGCACCATGATTAATTCAGGCAATCAAGGCTGCCCTCAGAAAGT 6444
OY 3725 CTCGAGCTCTACATCTCTCACTTATCATCATGTATAGTCTTGTGGAGAGATGGCA 3784
Db 6445 CTCGAGCTCTACATCTCTCACTTATCATCATGTATAGTCTTGTGGAGAGATGGCA 6504
OY 3785 GACTATGAGGAAATTCACATGGAACCTTAATGCTCTTGTGGCAATGTCGATTCATC 3844
Db 6505 GACTATGAGGAAATTCACATGGAACCTTAATGCTCTTGTGGCAATGTCGATTCATC 6564
OY 3845 TGGATTAACCAATATTTTAACTTCAATTTATGCTGATACATCCGTTTGACCC 3904
Db 6565 TGGATTAACCAATATTTTAACTTCAATTTATGCTGATACATCCGTTTGACCC 6624

QY 3905 AACATATTATAGCATTCGACGACTCTTCGATGGAGTGTAGTGGCTGTGATTTAAATAG 3964
 |||||||
 Db 6625 AACATATTATAGCATTCGACGACTCTTCGATGGAGTGTAGTGGCTGTGATTTAAATAG 6664
 |||||||
 QY 3965 TTGCAGCATGCCATTGGGAATGGAGAGTAAGCAATATCAGATGCAGATTACTGCTTC 4024
 |||||||
 Db 6685 TTGCAGCATGCCATTGGGAATGGAGAGTAAGCAATATCAGATGCAGATTACTGCTTC 6744
 |||||||
 QY 4025 ATCCACTTTTACCAATATGTTTGGCCACTGGTCTCTTCAAAAAGTGCATCTCACTTCA 4084
 |||||||
 Db 6745 ATCCACTTTTACCAATATGTTTGGCCACTGGTCTCTTCAAAAAGTGCATCTCACTTCA 6804
 |||||||
 QY 4085 AGGAGAGTAATGCTGAGAGTCCAGTCAATATCAAAAAGTGGCTGCAAGTGA 4144
 |||||||
 Db 6805 AGGAGAGTAATGCTGAGAGTCCAGTCAATATCAAAAAGTGGCTGCAAGTGA 6864
 |||||||
 QY 4145 CTTCCAGAGACAAATGAAGTACAGAGAGTAAGTACTCAGGAGTAATAATCTGCTTAC 4204
 |||||||
 Db 6865 CTTCCAGAGACAAATGAAGTACAGAGAGTAAGTACTCAGGAGTAATAATCTGCTTAC 6924
 |||||||
 QY 4205 CAGCATGATGTGAAGAGTCCATCTCAGACAGTCAATATGCCATCTGACTCT 4264
 |||||||
 Db 6925 CAGCATGATGTGAAGAGTCCATCTCAGACAGTCAATATGCCATCTGACTCT 6984
 |||||||
 QY 4265 CTTTTCAGAAATGCAAAAGTAAAGTCTTTCAGGAAATCAAGACTCTTCAACCTGT 4324
 |||||||
 Db 6985 CTTTTCAGAAATGCAAAAGTAAAGTCTTTCAGGAAATCAAGACTCTTCAACCTGT 7044
 |||||||
 QY 4325 GGTAACTCTTACAGCCAGCTTACTGACTGCTTACCTTGAATTCACCCAGAGTTG 4384
 |||||||
 Db 7045 GGTAACTCTTACAGCCAGCTTACTGACTGCTTACCTTGAATTCACCCAGAGTTG 7104
 |||||||
 QY 4385 GGTGACACAGATTCCTCTGAGAGTGAAGTCTGGGCTGGAGGACAGAGACCTTACTG 4444
 |||||||
 Db 7105 GGTGACACAGATTCCTCTGAGAGTGAAGTCTGGGCTGGAGGACAGAGACCTTACTG 7164
 |||||||
 QY 4445 AGGTTGCGCACTGACAGACCTGCGCTGACCTCTCTCTCTCAGCTCCAGGGCAG 4504
 |||||||
 Db 7165 AGGTTGCGCACTGACAGACCTGCGCTGACCTCTCTCTCTCAGCTCCAGGGCAG 7224
 |||||||
 QY 4505 TGTCCTCCCTGGCTGCTTCTTACCTTGTGCTAAATCTCTAGACAGACTGCTTGAAG 4564
 |||||||
 Db 7225 TGTCCTCCCTGGCTGCTTCTTACCTTGTGCTAAATCTCTAGACAGACTGCTTGAAG 7284
 |||||||
 QY 4565 CTTCTGAAATTAATATCATCATGCTGCTGCTTCTTGTGGGGGGCCAGAGAGGTCAT 4624
 |||||||
 Db 7285 CTTCTGAAATTAATATCATCATGCTGCTGCTTCTTGTGGGGGGCCAGAGAGGTCAT 7344
 |||||||
 QY 4625 CCAATTTAACTTAATCTTAATCTTCTGCAAGCTGCTCCAGATTAATCTCTCTTCC 4684
 |||||||
 Db 7345 CCAATTTAACTTAATCTTCTTAATCTTCTGCAAGCTGCTCCAGATTAATCTCTCTTCC 7404
 |||||||
 QY 4685 AATATACTAGGCAAAAGAGAGAGAGAACTGATGAAGCAATCTTCTCTGAAA 4744
 |||||||
 Db 7405 AATATACTAGGCAAAAGAGAGAGAGAACTGATGAAGCAATCTTCTCTGAAA 7464
 |||||||
 QY 4745 GTTAGGCTCTCAGAGTCAACACTTCTCTGTTGTAGAAAAAATATGATGAACCTTG 4804
 |||||||
 Db 7465 GTTAGGCTCTCAGAGTCAACACTTCTCTGTTGTAGAAAAAATATGATGAACCTTG 7524
 |||||||
 QY 4805 AAAAAGATATTTATGATGT 4824
 |||||||
 Db 7525 AAAAAGATATTTATGATGT 7544
 |||||||

RESULT 10
 HUMFVII
 LOCUS HUMFVII 9009 bp mRNA PRI 08-NOV-1994
 DEFINITION Human coagulation factor VIII:C (anti-hemophilic factor) mRNA.
 ACCESSION K01740
 NID 9182802
 VERSION K01740.1 GI:182802
 KEYWORDS factor VIII; hemophilia.

SOURCE Human cDNA to liver mRNA, and genomic DNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 9009)
 AUTHORS Toole,J.J., Knopf,J.L., Wozney,J.M., Sultzman,L.A., Becker,J.L.,
 Pittman,D.D., Kaufman,R.J., Brown,E., Shoemaker,C., Orr,E.C.,
 Amphlett,G.W., Foster,W.B., Coe,M.L., Knutson,G.J., Fass,D.N. and
 Hewick,R.M.
 TITLE Molecular cloning of a cDNA encoding human antihemophilic factor
 JOURNAL Nature 312 (5992), 342-347 (1984)
 COMMENT 85061550
 FEATURES
 source
 gene
 CDS

Location/Qualifiers
 1..9009
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="Xq28"
 151..7206
 /gene="F8C"
 151..7206
 /gene="F8C"
 /note="factor VIII"
 /codon_start=1
 /db_xref="GB:500-119-124"
 /protein_id="AA52484.1"
 /db_xref="pid:g182803"
 /db_xref="gi:182803"
 /translation="MDELSTCFELCLRFCSATRRYLIGAVELSDMYDGLGELP
 VDARPPVPKSPFNTSVYKTLFEYFVHLNIAAPRPMMGLPTLOAEVYD
 VITLKNASHPSVLAHGVSYKSLSEAEADVDODSOEKEEDVPFGSGSTYVOWL
 KENGPMSADPLCTLSYSHVDYKDNLSGLIGALVYREGSLAKETQTLAKTILF
 AVFDEGKMSHSETKNSLMODRDAASARAPMIGALVYNGVNSLPGILGCHKSYWHY
 IGMGTPPVHSLFLEGTFLVLRNRQSLSTPLTFPLTLMLODFLCHISLH
 OHGDEAVAVKVDSCPEEPOLMKNNNEAEDDDLDJSEMVDVFERDDNSPFLI
 VAKRKPWVHYIAAEEEDMDYAPLVLPDRSKSYQLNGPQIRKRYKVFMY
 IDTEFKTRALIOHSGILPLLYGEVGTPLTIFKNSARPNYNTIPGIDIVRLYSR
 RLPGVYKRLKDEPILPGELIPYKMYVYEOSPTSDPCLRTYSSVNMERDLASL
 IGPLICRVSVDKRGNOIMSDKRVILFSPEDNRKRYLTENLORLPNAGVOLD
 PEPQASNMHSINGVDSLOSLYCLHEVAWYTLISGAQDFLVSFGTTEHAKY
 YEDTLTLPESGETVFMSENPGLMILGHNSDRNMGMLALVSSCDKATGYED
 SYEDISATLILKNNALIEPRFSQNSRHSSTROKOPNATTPENDIEKTPFAHRTM
 PKIONVSSDLMLRLROSPYHGLSLDLOAKYETSDPSPAIDSNNSISEMTHP
 RPOLHSGDWETPESGLRLINKLGTATATLTKIDPFRVSSNNLSITPISMA
 AGTNTSLGPPSPMVHYDSDLODTLIGKSSPTBSGGPLSISEENNDKILBESMA
 NSOSSNKNVSSPESGLRFGRAHGPALITKNALEKYSILKTKNSNNSATNR
 KTHIDGRLIENSPTWONLIESDTEFKVTPLIHDMIMDKNATLRUNHMSNKT
 SSKNMEVQAKKEPILPDADNPDMSEFKMLFEPESARWLTORTGKNSLSCQPSPK
 QLVSLGPKSEVEGNFLSEKNKVVYGGEEFKDGLKEMVPSRNFLNLDLHN
 NTHNEKRIOEIEKEKTELIOENVVLOIHVGTGKNMNFILSTROMVEGSEGA
 YAPVLOPRLSNDSTNRKKTAFHRSKKGEEENEGTGNOKTOVEYACTRISPT
 SOONVTPQRSKRALKORPLPELELEKRIIVDTSTQSKNKHLPSTLQDNLVE
 KEKATIOSPLSDTLTSHSTIPQANRSPPLIAKVSFSPRIPLYLTFLVLODNSHL
 AASTRKDSGVQESSHFLOAKKNNLSIALITLMTDQDEBVGSLGTSANSVYKAY
 ENTVLPKPDLPKISGVELLPKVHLYOKDLPETSSNGSGHDLVGLSLOGEGEA
 KMNENANRGPKEPRLVATSESSAKPSKLLPLIADNMYGTOIPPEKESOEKSPK
 FKFKDITLSLNACESNAHAIANGOKPKELEVYMAOGRETCSONPVLKROKRE
 ITRTTLOSDEIDYDTISVEKKEEDPIYDDENSPSPFOKTHYTLAVERLM
 DYGMSSSPHYLRNAGSGVPOKRVYFOETDSFTQPLRYRGLNHLDLGLGYTA
 EVEDNIWVTEPRNOASRPYSFSSLSIYEEDROGAERKKNFVPRKRYTFWKVOHMA
 APTDEDFDCAWAFSDVDLEKDVHSLIGPLVCHNTLNPAHROVTVQDFLFT
 ILETKSMYFTENNERNCRAPCNIDQMDPEFKERYRPHDANGVYIMDLPGLVAAODR
 IRLWTLMSGNSNENHSHFSGHYEVYKKEPEKEMALYNLTGAEVTEYEMPSKAGTR
 VECLIGHLAAGNSHTFLVYSNKCOPLGMAASHIRPOITPAGSOYQAPKARLRY
 SCSINASTKEPESMIVDILAPVITIGITOGARQRFSSLYISQFTIMSLDGKMO
 TYRNSGTGLMVFEGVNDSSGILHNINPPIIARYLRLPHYSIRSLTLMELMGCDL
 NSCSMPLGMSKSAISDAQIYASSYFTNMFATWPSKARLHLQGRSNAMRQVNNPKM
 LOYDFQKTRVGTGTVGKSLTSMYKFEFLSSSDGGHOMTLFQNGKVKVFGQNG
 DSFTPVVNSLDPPLTLYRLRHQSWHQAALMAYLGCBAQDLV"

D	6066	CCATTCATTTCATTTCAGTGGAGATGTGTTCACTGACGAAAAAAGAGAGATTAATAAT	6125
Q	3365	GGCACTGTACAACTCTATATCCAGAGTGTGTTTGTAGACAGTGAATAGTTACATCCAAAGC	3424
D	6126	GGCAGCTACAACTCTATATCCAGAGTGTGTTTGTAGACAGTGAATAGTTACATCCAAAGC	6185
Q	3425	TGGAATTTGGCGGGGTGGAAATGCTTATTTGGCGAGCATTCATAGCTGGGATGAGACACT	3484
D	6186	TGGAATTTGGCGGGGTGGAAATGCTTATTTGGCGAGCATTCATAGCTGGGATGAGACACT	6245
Q	3485	TTTTCTGTGTAAAGCAATTAAGTGTCAAGCTCCCTGGGAAATGCTTCTGGACACATTAG	3544
D	6246	TTTTCTGTGTAAAGCAATTAAGTGTCAAGCTCCCTGGGAAATGCTTCTGGACACATTAG	6305
Q	3545	AGATTTTCAGATTATACAGCTTCAGAGCAATATGAGACGTGGGCGCCAAAGCTGGCCAGCT	3604
D	6306	AGATTTTCAGATTATACAGCTTCAGAGCAATATGAGACGTGGGCGCCAAAGCTGGCCAGCT	6365
Q	3605	TCATATATCCGGATCAATCAATGCTTGAGACCAAGAGACCCCTTTCTTGATCAAGT	3664
D	6366	TCATATATCCGGATCAATCAATGCTTGAGACCAAGAGACCCCTTTCTTGATCAAGT	6425
Q	3665	GGATCTGTGGCACCAATGATTTATACGGGATCAAGACCCAGGCTGCCCTGCAGAACTT	3722
D	6426	GGATCTGTGGCACCAATGATTTATACGGGATCAAGACCCAGGCTGCCCTGCAGAACTT	6485
Q	3725	CTCCAGGCTCTACATGCTCTCAGATTATCAATGATGTACTGTTGATGATGGGAAGAACTGGCA	3784
D	6486	CTCCAGGCTCTACATGCTCTCAGATTATCAATGATGTACTGTTGATGATGGGAAGAACTGGCA	6545
Q	3785	GACTTATCGAGAAATTCACATGGAACCTTAATGCTCTCTTTGGCAATGTGTGATTCATC	3844
D	6546	GACTTATCGAGAAATTCACATGGAACCTTAATGCTCTCTTTGGCAATGTGTGATTCATC	6605
Q	3845	TGGGATTAACAAACAAATTTTTTAACCTCCCAATATGTCTGATACATCCGTTTCACACC	3904
D	6606	TGGGATTAACAAACAAATTTTTTAACCTCCCAATATGTCTGATACATCCGTTTCACACC	6665
Q	3905	AACCATTAATAGCATTCGCGACGACCTTGGCAATGGATGGATGGGCTGTGATTTAAATAG	3964
D	6666	AACCATTAATAGCATTCGCGACGACCTTGGCAATGGATGGATGGGCTGTGATTTAAATAG	6725
Q	3965	TTGCAGCATCCATTGGGATGGAGATGAAGTAAGCAATATCAGATGCGACAGATTACTGCTTC	4024
D	6726	TTGCAGCATCCATTGGGATGGAGATGAAGTAAGCAATATCAGATGCGACAGATTACTGCTTC	6785
Q	4025	ATCCTACTTACCAATATGTTTGCCACGCTGTCTCCTTCAAAAGCTCGACTTCACCTCCA	4084
D	6786	ATCCTACTTACCAATATGTTTGCCACGCTGTCTCCTTCAAAAGCTCGACTTCACCTCCA	6845
Q	4085	AGGAGAGAGTAATGCTCGAGAGACCTCAGGTGAATATCCAAAGAGTGGCTGCAAGTGGGA	4144
D	6846	AGGAGAGAGTAATGCTCGAGAGACCTCAGGTGAATATCCAAAGAGTGGCTGCAAGTGGGA	6905
Q	4145	CTTTCACAACAACATGAAATGTCACAGAGTAATCTACTCAGGAGTAATAATCTCTGCTTAC	4204
D	6906	CTTTCACAACAACATGAAATGTCACAGAGTAATCTACTCAGGAGTAATAATCTCTGCTTAC	6965
Q	4205	CAGCATGTATGTAAGAGAGTTCCTCATCTCCAGACATCAAGATGGCCATCACTGAGACTCT	4264
D	6966	CAGCATGTATGTAAGAGAGTTCCTCATCTCCAGACATCAAGATGGCCATCACTGAGACTCT	7025
Q	4265	CTTTTTTCAAAATAGCAAGTAAGGTTTTTTCAGGAAATCAAGACTCTTTCACACCTGT	4324
D	7026	CTTTTTTCAAAATAGCAAGTAAGGTTTTTTCAGGAAATCAAGACTCTTTCACACCTGT	7085
Q	4325	GGTAACATCTCTAGACCAACCGTTACTGATGCGCTCTCAATTCACACCCCAAGAGTTG	4384
D	7086	GGTAACATCTCTAGACCAACCGTTACTGATGCGCTCTCAATTCACACCCCAAGAGTTG	7145
Q	4385	GGTGACACAGATTTGCCCTGAGATGGAGGTTCTGGGCTCGAGGACACAGAACTTACTGTG	4444

Db	7146	GGGACACCAAGTTGGCCCTGAGGATGAGGCTTTGGGCGTCGACGACAGGACCTCTACTG	7205
OY	4445	AGGGTGGCACTGCAGCACTGGCCACATGGCCGTACACTCTCCCTCAGCTCCAGAGGAC	4504
Db	7206	AGGGTGGCACTGCAGCACTGGCCACATGGCCGTACACTCTCCCTCAGCTCCAGAGGAC	7265
OY	4505	TGTCCCTCCCGGGCTCCCTTTACCTTTGTGCTAAATCCTAGCAGACACTGCTTTGAAG	4564
Db	7266	TGTCCCTCCCGGGCTCCCTTTACCTTTGTGCTAAATCCTAGCAGACACTGCTTTGAAG	7325
OY	4565	CCTCCTGATTAATCACTATCATCAGTCGCACTTCTTTGGTGGGGGGGCGAGAGGGTGCAT	4624
Db	7326	CCTCCTGATTAATCACTATCATCAGTCGCACTTCTTTGGTGGGGGGGCGAGAGGGTGCAT	7385
OY	4625	CCAATTTACTTAACCTTTACCTTAATTTTGTGACGCTGCTCCAGATTACTCTCTCTCTCC	4684
Db	7386	CCAATTTACTTAACCTTTACCTTAATTTTGTGACGCTGCTCCAGATTACTCTCTCTCTCC	7445
OY	4685	AATTAATACCTAGGCAAAAAGAGGAGAGAAACCTGCATGAAAGCAATTTCTTCCCTGAAAA	4744
Db	7446	AATTAATACCTAGGCAAAAAGAGGAGAGAAACCTGCATGAAAGCAATTTCTTCCCTGAAAA	7505
OY	4745	GTTTAGGCGCTCAGAGTCACACACTCTCTCGTGTGTGAAAAAATCTATGATGAAACTTGG	4804
Db	7506	GTTTAGGCGCTCAGAGTCACACACTCTCTCGTGTGTGAAAAAATCTATGATGAAACTTGG	7565
OY	4805	AAAAAGATATTTATGATGTT 4824	
Db	7566	AAAAAGATATTTATGATGTT 7585	

RESULT 11	LOCUS	DEFINITION	ACCESSION	NID	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	COMMENT
	HUMFVIIIC	9029 bp mRNA	PRI	08-NOV-1994																				
	Human coagulation factor VIII:C mRNA, complete cds.																							
	M4113																							
	g182817																							
	M4113.1	GI:182817																						
	coagulation factor VIII:C																							
	Human kidney, cDNA to mRNA, clones pF8[100,102]; and cell line																							
	GM416 DNA, clone pF8-4-3.																							
	Homo sapiens																							
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;																							
	Eutheria; Primates; Catarrhini; Hominoidea; Homo.																							
	1 (bases 1 to 9029)																							
	Truett,M.A., Blacher,R.W., Burke,R.L., Caputi,D., Chu,C., Dina,D.																							
	Harbog,K., Kuo,C.H., Maslarz,F.R., Merryweather,J.P., Najarian,R.																							
	Reich,C., Potter,S.J., Puma,J., Quiroga,M., Rall,L.B., Randolph,P.																							
	Urdea,M.S., Valenzuela,P., Dahl,H.-H.M., Favalaro,J., Hansen,J.,																							
	Nordang,O., and Ezdan,M.																							
	Characterization of the polypeptide composition of human factor																							
	VIII:C and the nucleotide sequence and expression of the human																							
	kidney cDNA																							
	DNA 4 (5), 333-349 (1985)																							
	860811564																							
	2 (sites)																							
	Yousouflian,H., Wong,C., Aronis,S., Platokoukis,H., Kazazian,H.H.																							
	Jr. and Antonarakis,S.E.																							
	Moderately severe hemophilia A resulting from Glu----Gly																							
	substitution in exon 7 of the factor VIII gene																							
	Am. J. Hum. Genet. 42 (6), 867-871 (1988)		</																					

The mutation at position 1042 results in a change of Glu to Lys and the one at position 1043 in a change from Glu to Gly. The mutation at position 1042 produces a premature stop codon.

location/Qualifiers

1. .9029

FEATURES

source

mRNA	1. .79029	/note="F-VIII mRNA"
gene	172. .72227	/gene="F8C"
sig_peptide	172. .228	/gene="F8C"
CDS	172. .72227	/note="coagulation factor VIII:C signal peptide"

/note="preprocoagulation factor VIII:C"
 /codon_start=1
 /db_xref="GDB:600-119-124"
 /protein_id="AA552485.1"
 /db_xref="PID:g182818"
 /db_xref="GI:182818"
 /translation="MOLESTCFPLCLRFCSATRRYYLGAVELSPMDYMOSDGEHL
 VDAPRPAPKSEPFNTSVYKKTLEFTEHDLMFNIAIPRPMWMLGLGTLQAVQYD
 VYTLTKMASHVYSLHAGVSWYKASDEAGVDQDQSRKEDDKAFEGSSHTYVQVYI
 KENGPAISPLDITLSTYSHDYLVKDLSGLGALTYCBRESGLAKETQTLHKTLLT
 AYDEKSKSWSETKNSLMQDRDASAKRAMPKHTVNGVNSLPLGLICHRKSWYWH
 IGGITPEVHSIFLESHFTVLNHRQASLETPIITFLAQTLMDLGGTFLFCHIS
 QAKHMEATVYVDSCEEPEDQLRKNNNEEDVDLTDSEMDVFERDDNSGFLQIR
 VQHRKMTWYIAAEEBMDYAPLVLAAPDRYSQKXQYLNNGPOIGIKRKYVMEFA
 TLEKFTFRAIOHESGILGFLLYGEGGTLITFENQASRRYINPHITVPRVLS
 PLTEGPKHLKDEPIGELFEIKKTKMTVGTGDTGDCPLCLRYYSFPMERDLAGI
 IGLILCYESVDGONGMOISMGRVILLPSVDENRMYLLENQRELPINPAQVLE
 PEQASINMHSINGYVPSDLSVCLHEVATYILISIQADIFLSVPSCKNTQGYED
 YEDTLTLPFEGSEYIVMSBENMGLILCHNSDFENRMTALMISVSDCKNTQGYED
 STEDISAVILKMLNLIEPRFSQNSRHSPQKQOFNATIPENDIEKTDPMFARTHP
 PIQVWNSDDLMLNLROSTPEHGLSLSDLOAKETTESDDSPGALISNLSLMTHTF
 RQLHSGMGVTPESGILQTLNLEKLTGTAATELKLDEKFTYSNNLISITPDNLIA
 AQTUNSSLGPPSMVYHDSQDLDTLFFKSKSPLESGGPLISSENNDSKLLPESGL
 NQESMGVNSVSTESGRLPFGKRAHGAPLLKNAALPKVSIILKTKTKSNSATNR
 KIHGDSPLLINSPWNONLIESPTERKATPLIHLMDLMDKNAALRLMHSNKT
 SKNNEMVQOKKEGPIPPDAQPDMSKFFKMLPEPSAKWIDQRTGKNSLNSGQSPK
 QLVSGEPSEVGEQNFLSEKNVAVGKEEFLVDGLKMPVPSRNLPINLMDIHEN
 NTHQNEKIQEIEIEKREGLIQENVLPDIHTVGTIKKMKMLFLSTRONEGSYDGA
 VAPVADPFLSNLSDNSTNRKTKTAHFSKGEENELGQNKQIYEKATCTRTSPNT
 SOONTYQSRKALDKQFPLEETLEKRIIVDTSTQWNSKMLHPSTITQDYNVE
 KKGKQITQSPDLSDCTLRSHSIPQANRSLPIAKVSSPSIRPIYITLVIFQNSHLP
 AASRKGKSGVQESSHFLQAKKNNLTALITLLEMTQDQREVGLGTSATNSYTKAV
 ENTVLPKDLPTSKVELLPKRYHLYQDLPLETSSQSPHDLVGSLSLQTEGAL
 KKNENRPGKQFELVNAIESSAKTPSKLDELANDNHYGTQIPREMKSOEKSEKTA
 FKFKKQITLISLACSENHAIANINBQNGOPELTVMAWQGRERLCSQNPVLRKHQRE
 IRTTLQSDQDEIDYDTISYEAKKEDPIDEDENQSPRFQKTEVYFIAAVERLMA
 DGMSSSHVLNRAQSGSVPOFKKVVQDETDSFQPLRGELNHLGLDGYIAA
 EVEDINFTFNRQASRPYSFYSLLSVEEDROGAPBRKPNVKNREKTKYFWYQHHM
 APTKEPQCKAMAYSPDYDELKQVHSGILGFLVLCYHCHNTLNPAGHGOVQERLE
 IPEDEKSWYFTEBNERNCRAPCNIQMDPEPTKENYRHNALINGIMDLPGLVMAQDR
 IRTVILMSGSDENHSHIFSGHVFYTRKREKYEKNAALNLPGVEYTEVMEPSKGIWR
 VECLLIGELHAGMHSILFVSNCKOTCPGASGSIIPFOIAYSGOYQAMAPKLARIY
 GSIISAMSTGKPEFWIKQVDLAPMIHGIRKIQAGRKFSLSISFTLIMLQKGMQ
 TYRGSTSTLWFLFVGNVDSGIKKINFINPEPIIAYIRLHPHYSIRSTLIMLQKGMQ
 NCSGALIMESKASIDQAQITLSSYFTWMPMAWPSKSKRLHQSNSNMROYNPKRM
 LQYDQKTKMYKQVLTQOYKSLTSMYKVELISSDDGHQWTLFQNGKVVQFQNO
 DSFTFYVNSLDPPLLRILKRIHPQSWHQIALRREVLCEXQDLY"
 898..71224
 /gene="F8C"
 /note="coagulation factor VIII:C"
 1042
 /gene="F8C"
 /note="g in wt; a in hemophiliac [2]"
 1043
 /gene="F8C"
 /note="a in wt; g in hemophiliac [2]"

mutation 6853
/gene="F8c"
/note="c in wt; t in hemophilic [3]"
BASE COUNT 2860 a 1910 c 1848 g 2411 t
ORIGIN 185 bp upstream of SacI site.

Query Match 51.1%; Score 2468.8; DB 9; Length 9029;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2473; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 2345 AAGAGCTTCGCCAGAACCCACAGCTTTGAAGCCCATCAAGCGGAATTAACCTGAC 2404
DB 5127 AAGGCTGCTCTCAAAACCCACAGCTTTGAAGCCCATCAAGCGGAATTAACCTGAC 5186
QY 2405 TACTCTCAGTCAGATCAAGAGAAATTAATGATGATACATATCAGTTGAATGAA 2464
DB 5187 TACTCTCAGTCAGATCAAGAGAAATTAATGATGATACATATCAGTTGAATGAA 5246
QY 2465 GAAGGAAGATTTGACATTTATGATGAGATGAATATCAGCCCGGACCTTCAAAA 2524
DB 5247 GAAGGAAGATTTGACATTTATGATGAGATGAATATCAGCCCGGACCTTCAAAA 5306
QY 2525 GAAACACAGACATTTATTTATGCTGAGTGAGAGGCTGGGATTAAGGATGAGTAG 2584
DB 5307 GAAACACAGACATTTATTTATGCTGAGTGAGAGGCTGGGATTAAGGATGAGTAG 5366
QY 2585 CTCGCCACATGTTCTAAGAAACAGGGCTCAGATGGCAGTGCTCCCTCAGTTCAAGAAAGT 2644
DB 5367 CTCGCCACATGTTCTAAGAAACAGGGCTCAGATGGCAGTGCTCCCTCAGTTCAAGAAAGT 5426
QY 2645 TGTTCCTCAGGAATTTACTGATGCTCCTTACTACAGCCCTTATACGTTGGAAGACTAAA 2704
DB 5427 TGTTCCTCAGGAATTTACTGATGCTCCTTACTACAGCCCTTATACGTTGGAAGACTAAA 5486
QY 2705 TGAACATTTGGGACTCTGGGGCCATATATAGAGCAGAGAAATTAAGATTAATCATGCT 2764
DB 5487 TGAACATTTGGGACTCTGGGGCCATATATAGAGCAGAGAAATTAAGATTAATCATGCT 5546
QY 2765 AACTTTGAGAAATCAGGCTCTGCTCCCTATCTCTTATCTAGCCCTTATTTTATGA 2824
DB 5547 AACTTTGAGAAATCAGGCTCTGCTCCCTATCTCTTATCTAGCCCTTATTTTATGA 5606
QY 2825 GGAAGATCAGAGGCAAGAGAGAGAACTAGAAAACTTTGTCAAGCTTATGAAGCAA 2884
DB 5607 GGAAGATCAGAGGCAAGAGAGAGAACTAGAAAACTTTGTCAAGCTTATGAAGCAA 5666
QY 2885 AACTTACTTTGGAAGAGCAACATCATATGACCCCACTAAAGATGATGATGATGCA 2944
DB 5667 AACTTACTTTGGAAGAGCAACATCATATGACCCCACTAAAGATGATGATGATGCA 5726
QY 2945 AGCTGGGCTTATTTCTGATGATGATGATGATGATGATGATGATGATGATGATG 3004
DB 5727 AGCTGGGCTTATTTCTGATGATGATGATGATGATGATGATGATGATGATGATG 5786
QY 3005 ACCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3064
DB 5787 ACCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5846
QY 3065 ACAGGAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3124
DB 5847 ACAGGAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5906
QY 3125 AAATATGGAAGAAAGCTGAGGGCTCCCTGCAATATCCAGATGGAAGATCCACATTTAA 3184
DB 5907 AAATATGGAAGAAAGCTGAGGGCTCCCTGCAATATCCAGATGGAAGATCCACATTTAA 5966
QY 3185 AGAGATTAATGCTTCCATGCAATCAATGCTTACATTAATGATGATGATGATGATGATG 3244
DB 5967 AGAGATTAATGCTTCCATGCAATCAATGCTTACATTAATGATGATGATGATGATGATG 6026
QY 3245 AATGCTCAGATCAAGAGATTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3304
DB 3304 AATGCTCAGATCAAGAGATTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT

DB 6027 AATGCTCAGATCAAGAGATTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6086
QY 3305 CCATTTATTCATTTTCACTGAGACATGCTTCACTGTAACGAAAAAGAGAGATTAAT 3364
DB 6087 CCATTTATTCATTTTCACTGAGACATGCTTCACTGTAACGAAAAAGAGAGATTAAT 6146
QY 3365 GGCACGTACAACTCTATCCAGGCTGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 3424
DB 6147 GGCACGTACAACTCTATCCAGGCTGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 6206
QY 3425 TGAATTTGGCGGGTGAAGATCCCTTATGCGAGACATCACTGATGATGATGATGATG 3484
DB 6207 TGAATTTGGCGGGTGAAGATCCCTTATGCGAGACATCACTGATGATGATGATGATG 6266
QY 3485 TTTTCTGCTGACAGCAATTAAGTCTCAGACTCCCTGGGAAATGCTTCTGACATTTAG 3544
DB 6267 TTTTCTGCTGACAGCAATTAAGTCTCAGACTCCCTGGGAAATGCTTCTGACATTTAG 6326
QY 3545 AGATTTTCAAGTTACAGCTTCAGAGCAATTAAGAGAGAGAGAGAGAGAGAGAGAG 3604
DB 6327 AGATTTTCAAGTTACAGCTTCAGAGCAATTAAGAGAGAGAGAGAGAGAGAGAGAG 6386
QY 3605 TCATTTATCCGAGATCAATCAATGCTGAGACACCAAGAGAGAGAGAGAGAGAGAGAG 3664
DB 6387 TCATTTATCCGAGATCAATCAATGCTGAGACACCAAGAGAGAGAGAGAGAGAGAG 6446
QY 3665 GGATCTGTTGGCACCAATGATTAATTCACGGCATCAAGACCAAGAGAGAGAGAGAG 3724
DB 6447 GGATCTGTTGGCACCAATGATTAATTCACGGCATCAAGACCAAGAGAGAGAGAG 6506
QY 3725 CTCAGGCTCTACATCTCTCAGTTATCATCATCATCATCATCATCATCATCATCAT 3784
DB 6507 CTCAGGCTCTACATCTCTCAGTTATCATCATCATCATCATCATCATCATCATCAT 6566
QY 3785 GACTTATCAGAGAAATTCACCTGGAACCTTAATGCTTCTTGGCAATGATGATGATG 3844
DB 6567 GACTTATCAGAGAAATTCACCTGGAACCTTAATGCTTCTTGGCAATGATGATGATG 6626
QY 3845 TGGGATTAACCAATATTTTAAACCTCCCAATTAATGCTGATGATGATGATGATG 3904
DB 6627 TGGGATTAACCAATATTTTAAACCTCCCAATTAATGCTGATGATGATGATGATG 6686
QY 3905 AACTCATTAATGATTTGGGAGACCTCTTCCGATGATGATGATGATGATGATGATG 3964
DB 6687 AACTCATTAATGATTTGGGAGACCTCTTCCGATGATGATGATGATGATGATGATG 6746
QY 3965 TTGAGCATGCTTGGGAAATGAGAGTAAGCAATATCAGATGATGATGATGATGATG 4024
DB 6747 TTGAGCATGCTTGGGAAATGAGAGTAAGCAATATCAGATGATGATGATGATGATG 6806
QY 4025 ATCTTACTTTAACAATATGTTTGGCACCTGCTCTTCAAAAAGCTGACTTCACTCA 4084
DB 6807 ATCTTACTTTAACAATATGTTTGGCACCTGCTCTTCAAAAAGCTGACTTCACTCA 6866
QY 4085 AGGAGAGATTAATCCCTGAGAGACCTCAGTGATTAATTAATTAATTAATTAATTA 4144
DB 6867 AGGAGAGATTAATCCCTGAGAGACCTCAGTGATTAATTAATTAATTAATTAATTA 6206
QY 4145 CTTCAGAAAGCAATGAAAGTCAAGAGATTAATTAATTAATTAATTAATTAATTA 4204
DB 6927 CTTCAGAAAGCAATGAAAGTCAAGAGATTAATTAATTAATTAATTAATTAATTA 6986
QY 4205 CAGCATGATGATGAGAGATTTCTCATCTCCAGAGTCAAGAGTCAAGAGTCAAG 4264
DB 6987 CAGCATGATGATGAGAGATTTCTCATCTCCAGAGTCAAGAGTCAAGAGTCAAG 7046
QY 4265 CTTTTCAGAAATGAGCAAGATTTTTCAGAGAAATCAAACTCTTCAACACTGT 4324
DB 7047 CTTTTCAGAAATGAGCAAGATTTTTCAGAGAAATCAAACTCTTCAACACTGT 7106
QY 4325 GGTGAACCTCTAGACCAAGCTTACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 4384
DB 7107 GGTGAACCTCTAGACCAAGCTTACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 7166

QY	2645	TGTTTTCCAGGAATTTACTGATGGCTCTTTTACTCAGCCCTTATACCGTCGAGAACTAAA	2704
Dp	5365	TGTTTTCCAGGAATTTACTGATGGCTCTTTTACTCAGCCCTTATACCGTCGAGAACTAAA	5424
QY	2705	TGAACATTTGGGACCTCTGGGGCCATATATAGAGCAGAAAGTTGAAGTAATATCATGCT	2764
Dp	5425	TGAACATTTGGGACCTCTGGGGCCATATATAGAGCAGAAAGTTGAAGTAATATCATGCT	5484
QY	2765	AACTTTCAGAAATCAGGCTCTCGTCCATATTCCTCTATTTACTACCTTATTTCTTATGA	2834
Dp	5485	AACTTTCAGAAATCAGGCTCTCGTCCATATTCCTCTATTTACTACCTTATTTCTTATGA	5544
QY	2835	GGAAGATCAGAGCCAGGACAGACCACTAGAAAAAACTTTGTCAAAGCTTAATGAACCAA	2884
Dp	5545	GGAAGATCAGAGCCAGGACAGACCACTAGAAAAAACTTTGTCAAAGCTTAATGAACCAA	5604
QY	2885	AACTTACTTTTGGAAAGTGCACATATATATAGCAACCCATTAAGATGAGTTTGACTGAA	2944
Dp	5605	AACTTACTTTTGGAAAGTGCACATATATATAGCAACCCATTAAGATGAGTTTGACTGAA	5664
QY	2945	AGCCTGGGCTTATTTCTCATGTTGACCTGCGAAAAAGATGTCACCTCAGGCTCATGG	3004
Dp	5665	AGCCTGGGCTTATTTCTCATGTTGACCTGCGAAAAAGATGTCACCTCAGGCTCATGG	5724
QY	3005	ACCCTTCTGGTCTGCCACACTAACACACTGAACCTGCTCATGGGAGACAAGTACAGT	3064
Dp	5725	ACCCTTCTGGTCTGCCACACTAACACACTGAACCTGCTCATGGGAGACAAGTACAGT	5784
QY	3065	ACAGGAATTTGCTCTGTTTTTTCACCATCTTTTATATGAGACCAAAAGCTGGTACTTCACGA	3124
Dp	5785	ACAGGAATTTGCTCTGTTTTTTCACCATCTTTTATATGAGACCAAAAGCTGGTACTTCACGA	5844
QY	3125	AAATATGGAAGAAACTCGAGGGCTCCCTGTCATATCCAGATGAGAAATCCACTTTTAA	3184
Dp	5845	AAATATGGAAGAAACTCGAGGGCTCCCTGTCATATCCAGATGAGAAATCCACTTTTAA	5904
QY	3185	AGAGAAATTTATGCTTCCATGCAATTCATATGGCTATACATATATGATACACTACCTGGCTAGT	3244
Dp	5905	AGAGAAATTTATGCTTCCATGCAATTCATATGGCTATACATATATGATACACTACCTGGCTAGT	5964
QY	3245	AATGGCTCAGGATCAAAGAAATTTGATGATGATCTGCTCAGACATGGGCACGAATGAAAAAT	3304
Dp	5965	AATGGCTCAGGATCAAAGAAATTTGATGATGATCTGCTCAGACATGGGCACGAATGAAAAAT	6024
QY	3305	CCATTCTATTTCATTTCTAGTGGACATGTTGCTACCTGTACGAAAAAAAGAGAGTATTAAT	3364
Dp	6025	CCATTCTATTTCATTTCTAGTGGACATGTTGCTACCTGTACGAAAAAAAGAGAGTATTAAT	6084
QY	3365	GGCACTGTACAAATCTATATCCAGGCTGTTTTTGTAGACAGTGGAAATGTTACATCCAAAGC	3424
Dp	6085	GGCACTGTACAAATCTATATCCAGGCTGTTTTTGTAGACAGTGGAAATGTTACATCCAAAGC	6144
QY	3425	TGGAATTTGGCGGGTGGAAATGCTTATTTGGCGAGCATCTACATGCTGGGATGAGCAGCT	3484
Dp	6145	TGGAATTTGGCGGGTGGAAATGCTTATTTGGCGAGCATCTACATGCTGGGATGAGCAGCT	6204
QY	3485	TTTTCTGGTGTACGCAATTAAGTGTACAGATCTCCCTGGGAAATGGGCTTCTGGACACATTAG	3544
Dp	6205	TTTTCTGGTGTACGCAATTAAGTGTACAGATCTCCCTGGGAAATGGGCTTCTGGACACATTAG	6264
QY	3545	AGATTTTCAGATTTACACCTTCAGGACAAATATGTGACAGTGGGCCCAAAAGCTGGCAGACT	3604
Dp	6265	AGATTTTCAGATTTACACCTTCAGGACAAATATGTGACAGTGGGCCCAAAAGCTGGCAGACT	6324
QY	3605	TCATTTATTTCCGGATCAATCAATGCTGTGAGACACCAAGAGAGCCCTTTTCTTGGATCAAGT	3664
Dp	6325	TCATTTATTTCCGGATCAATCAATGCTGTGAGACACCAAGAGAGCCCTTTTCTTGGATCAAGT	6384
QY	3665	GGATCTGTGGCACCAATGATTAATTCACGGGCAATCAAGACCCAGGGGGCCCGTCACAAATTT	3724
Dp	6385	GGATCTGTGGCACCAATGATTAATTCACGGGCAATCAAGACCCAGGGGGCCCGTCACAAATTT	6444
QY	3725	CTCCAGCCTCTACATCTCTCAGTTTATCATCATGATGATAGTCTTGATGGGAGAAAGTGGCA	3784

Db	6445	CTCCAGCCTCTACACTCTCTAGTTTATCATCATGTAATACTCTTGATGGGAAGAACTGGCA		6504
QY	3785	GACATTATCGAGGAAATTCACCTGGAAACCTTAATAGCTCTTTGGCAATGAGATTCAC		3844
Db	6505	GACTTATCGAGGAAATTCACCTGGAAACCTTAATAGCTCTTTGGCAATGAGATTCAC		6564
QY	3845	TGGGATTAACACACAAATATTTTATACCCCTCCCAATATATGCTGCATACATCCGTTGCACC		3904
Db	6565	TGGGATTAACACACAAATATTTTATACCCCTCCCAATATATGCTGCATACATCCGTTGCACC		6624
QY	3905	AACATCATTAATAGCACTTCGACACACTCTTCGCATGAGTGGCTGTGATTTAAATAG		3964
Db	6625	AACATCATTAATAGCACTTCGACACACTCTTCGCATGAGTGGCTGTGATTTAAATAG		6684
QY	3965	TTTGACATATGCATTTGGGAAATGGAGATGAAGCAATATAGATAGCAAGATTAACGCTTC		4024
Db	6685	TTTGACATATGCATTTGGGAAATGGAGATGAAGCAATATAGATAGCAAGATTAACGCTTC		6744
QY	4025	ATTCCTACTTTACCAATATATGTTTTCACACCTGCTCTCTTCAAAAAGCTCGACTTCACCTCA		4084
Db	6745	ATTCCTACTTTACCAATATATGTTTTCACACCTGCTCTCTTCAAAAAGCTCGACTTCACCTCA		6804
QY	4085	AGGGAGAGTAATGCTCTGGAGACCTCAGGTGAATTAATCCAAAAGAGTGGCTGCAAGTGA		4144
Db	6805	AGGGAGAGTAATGCTCTGGAGACCTCAGGTGAATTAATCCAAAAGAGTGGCTGCAAGTGA		6864
QY	4145	CTTCCAGAAAGCAATGAAATATCAGAGAGTAATACTACAGGAGATTAATCTCTCTTAC		4204
Db	6865	CTTCCAGAAAGCAATGAAATATCAGAGAGTAATACTACAGGAGATTAATCTCTCTTAC		6924
QY	4205	CAGCATGATATGAAAGAGATTTCTCATCTTCACACACTCAAGATGGCATCAGTGCATCT		4264
Db	6925	CAGCATGATATGAAAGAGATTTCTCATCTTCACACACTCAAGATGGCATCAGTGCATCT		6984
QY	4265	CTTTTTCAGAAATGCAAGTAAAGGTTTTTCAGGAAATCAAGACCTCTTACACCTGT		4324
Db	6985	CTTTTTCAGAAATGCAAGTAAAGGTTTTTCAGGAAATCAAGACCTCTTACACCTGT		7044
QY	4325	GSTGAACCTCTTAACCCACCGTTACTGATGATCGCTACCTTCCAAATTCACCCCAAGATG		4384
Db	7045	GSTGAACCTCTTAACCCACCGTTACTGATGATCGCTACCTTCCAAATTCACCCCAAGATG		7104
QY	4385	GGTGCACCAAGATTCCTCTGAGAGATGAGAGTTCGGGCTCGAGAGCACAGACACTCTACTG		4444
Db	7105	GGTGCACCAAGATTCCTCTGAGAGATGAGAGTTCGGGCTCGAGAGCACAGACACTCTACTG		7164
QY	4445	AGGGTGGCACTGCAGACACTGACACTGCGCTACACTTCCTCTCAGCTCCAGGGCAG		4504
Db	7165	AGGGTGGCACTGCAGACACTGACACTGCGCTACACTTCCTCTCAGCTCCAGGGCAG		7224
QY	4505	TGTCCCTCCCTGCGCTTGCCTTCTACCTTGTGCTAAATCCTTAGACACACTGCTTGAAG		4564
Db	7225	TGTCCCTCCCTGCGCTTGCCTTCTACCTTGTGCTAAATCCTTAGACACACTGCTTGAAG		7284
QY	4565	CTCTCGAATTAACATATACATCAGTCGCACTTCTTGGTGGGGGGCCAGAGGGTGCAT		4624
Db	7285	CTCTCGAATTAACATATACATCAGTCGCACTTCTTGGTGGGGGGCCAGAGGGTGCAT		7344
QY	4625	CCAATTTAACTTAACCTTACTACTAATTTTTCGACAGTGTCCAGATTAATCTCTCTTCC		4684
Db	7345	CCAATTTAACTTAACCTTACTACTAATTTTTCGACAGTGTCCAGATTAATCTCTCTTCC		7404
QY	4685	AATATTAAGTAGGCAAAAAGAGTAGAGAGAAACCTGCATGAAAGCAATTCCTCCGTAAGA		4744
Db	7405	AATATTAAGTAGGCAAAAAGAGTAGAGAGAAACCTGCATGAAAGCAATTCCTCCGTAAGA		7464
QY	4745	GTTAGGCTCTCAGAGTCACACACTCTCTGTTGTGTAAGAAAACATATGATGAATAACTTG		4804
Db	7465	GTTAGGCTCTCAGAGTCACACACTCTCTGTTGTGTAAGAAAACATATGATGAATAACTTG		7524
QY	4805	AAAAAGATTTATGATGTT 4824		


```

Db 522 GTCAAGGAGAGAGATGATGAAGTCTTCCTGGTGAAGCCATCATATGCTGGC 581
QY 544 AGGTCCTGAAGAGATGCTCAATGGCTGTGACCCACTGTGCTTACTACTCATATC 603
Db 582 AGGTCCTGAAGAGATGCTCAATGGCTGTGACCCACTGTGCTTACTACTCATATC 641
QY 604 TTTCTCATGTGAGCTGTGTAAGACTGTAATGAGGCTCATGAGGCTTACTAGTAT 663
Db 642 TTTCTCATGTGAGCTGTGTAAGACTGTAATGAGGCTCATGAGGCTTACTAGTAT 701
QY 664 GTAGAGAGAGAGTCTGGCCAGAGAAAAGACACAGACCTTGCACAAATTTACTACTT 723
Db 702 GTAGAGAGAGAGTCTGGCCAGAGAAAAGACACAGACCTTGCACAAATTTACTACTT 761
QY 724 TTTGCTGATTTGATGAGAGGAAAAGTTGGCACTAGAAAACAGAACTCTTGATGAGG 783
Db 762 TTTGCTGATTTGATGAGAGGAAAAGTTGGCACTAGAAAACAGAACTCTTGATGAGG 821
QY 784 ATAGAGATGCTGATCTGCTGGGCTGGGCTTAATGCAACAGTCAATGCTTAA 843
Db 822 ATAGAGATGCTGATCTGCTGGGCTGGGCTTAATGCAACAGTCAATGCTTAA 881
QY 844 ACAGGCTCTGCCAGTCTGATTTGGATGCCACAGAAATGATGCTTATGGCATGATG 903
Db 882 ACAGGCTCTGCCAGTCTGATTTGGATGCCACAGAAATGATGCTTATGGCATGATG 941
QY 904 GAATGGGACACACCTCCGAAAGTGCATCAATTTCTCGAAGGTGCACAACTTTCTGTGA 963
Db 942 GAATGGGACACACCTCCGAAAGTGCATCAATTTCTCGAAGGTGCACAACTTTCTGTGA 1001
QY 964 GGAACCATGCCAGGCGCTCTTGAATCTGCCAATTAATCTTCTTACAGCTCAAAAC 1023
Db 1002 GGAACCATGCCAGGCGCTCTTGAATCTGCCAATTAATCTTCTTACAGCTCAAAAC 1061
QY 1024 TCTTGTGATGACCTTGGACACTTCTTCTTGTGCTATATCTCTCCCAACATGATG 1083
Db 1062 TCTTGTGATGACCTTGGACACTTCTTCTTGTGCTATATCTCTCCCAACATGATG 1121
QY 1084 GCATGGAGCTTATGTAAGTAGACAGCTGTCCAGAGAAACCCCACTACGAATGAAA 1143
Db 1122 GCATGGAGCTTATGTAAGTAGACAGCTGTCCAGAGAAACCCCACTACGAATGAAA 1181
QY 1144 ATATGAGAGAGGAGAGACTATGATGATGATCTTACTGATTTGGAATGATGCTG 1203
Db 1182 ATATGAGAGAGGAGAGACTATGATGATGATCTTACTGATTTGGAATGATGCTG 1241
QY 1204 GGTTTGATGATGACACTTCTTCTTCTTATCCAAATTCGCTCAGTTGCCAAGACATC 1263
Db 1242 GGTTTGATGATGACACTTCTTCTTCTTATCCAAATTCGCTCAGTTGCCAAGACATC 1301
QY 1264 CTAAACCTTGGGTACATTTACATTTGCTGTAAGAGAGAGAGAGAGAGAGAGAGAG 1323
Db 1302 CTAAACCTTGGGTACATTTACATTTGCTGTAAGAGAGAGAGAGAGAGAGAGAGAG 1361
QY 1324 TCTTCCGCCCCGATGACAGAGATTATAAAAGTCAATATTTTGAACAATGGCCCTCAGCGA 1383
Db 1362 TCTTCCGCCCCGATGACAGAGATTATAAAAGTCAATATTTGAACAATGGCCCTCAGCGA 1421
QY 1384 TTTGCTGAGAGAGTACAAAAAGTCCGATTTATGSCATACACAGATGAACCTTTAAGACTC 1443
Db 1422 TTTGCTGAGAGAGTACAAAAAGTCCGATTTATGSCATACACAGATGAACCTTTAAGACTC 1481
QY 1444 GTGAAGCTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1503
Db 1482 GTGAAGCTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1541
QY 1504 ACACACTGTTGATTTATTTAAGAAATCAAGCAAGCAAGCAATTAACATCTACCTCAGC 1563
Db 1542 ACACACTGTTGATTTATTTAAGAAATCAAGCAAGCAAGCAATTAACATCTACCTCAGC 1601
QY 1564 GAATGACCTGATGCTGCTTGTGATTTCAAGAGATTAACCAAAAGGTGTAAAAATTTGA 1623
Db 1602 GAATGACCTGATGCTGCTTGTGATTTCAAGAGATTAACCAAAAGGTGTAAAAATTTGA 1661

```

```

QY 1624 AGGATTTTCCAAATTTGCCAGAGAAAATATTCAAATATATAATGACAGTACTAGAAG 1683
Db 1662 AGGATTTTCCAAATTTGCCAGAGAAAATATTCAAATATATAATGACAGTACTAGAAG 1721
QY 1684 ATGGGCCAACTAAATCAGATCTCGGTGCTGACCCGCTATTAATCTAGTTTGGTAAAT 1743
Db 1722 ATGGGCCAACTAAATCAGATCTCGGTGCTGACCCGCTATTAATCTAGTTTGGTAAAT 1781
QY 1744 TGGAGAGATCTAGGCTCAGAGCTCATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1803
Db 1782 TGGAGAGATCTAGGCTCAGAGCTCATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1841
QY 1804 TAGATCAAGAGAGAAACAGATATGTCAGACAGAGAAATGATCATCGTTTCTGAT 1863
Db 1842 TAGATCAAGAGAGAAACAGATATGTCAGACAGAGAAATGATCATCGTTTCTGAT 1901
QY 1864 TTGATGAGAACCGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1923
Db 1902 TTGATGAGAACCGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1961
QY 1924 CTGAGATGAGAGCTTGGAGATCCAGAGTTCCAGACCTCCCAACATCATGACATCATG 1983
Db 1962 CTGAGATGAGAGCTTGGAGATCCAGAGTTCCAGACCTCCCAACATCATGACATCATG 2021
QY 1984 GCTATGTTTGTATGATTTGTCAGTTGTCAGTTGTTGTTGTCATGAGTGGCATCTGATG 2043
Db 2022 GCTATGTTTGTATGATTTGTCAGTTGTCAGTTGTTGTTGTCATGAGTGGCATCTGATG 2081
QY 2044 TTCTAAGCATTTGAGACAGACTGACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCA 2103
Db 2082 TTCTAAGCATTTGAGACAGACTGACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCA 2141
QY 2104 AACACAAATGGCTATGATGAGACACACTCACTTCCATTTCCAGAGAAATGCT 2163
Db 2142 AACACAAATGGCTATGATGAGACACACTCACTTCCATTTCCAGAGAAATGCT 2201
QY 2164 TCATGTCATGAGAAAAACCCAGGCTATGATTTCTGGGGTGCACAACTCAGACTTTCGGA 2223
Db 2202 TCATGTCATGAGAAAAACCCAGGCTATGATTTCTGGGGTGCACAACTCAGACTTTCGGA 2261
QY 2224 ACAGAGGCAATGACCCGCTTACTGAAAGTTTCTAGTTTGACAGAACTGGTATAT 2283
Db 2262 ACAGAGGCAATGACCCGCTTACTGAAAGTTTCTAGTTTGACAGAACTGGTATAT 2321
QY 2284 ACAGAGGCAATGATGATGATTTTCTAGCATTTCTGAGTAAATGATGATGATGATGATG 2343
Db 2322 ACAGAGGCAATGATGATGATTTTCTAGCATTTCTGAGTAAATGATGATGATGATGATG 2381
QY 2344 CAAGAAGCTTCTCCAGAAACCCAGGCTTGTGAAAGCCATCAACGGGAAATTAATCGTA 2403
Db 2382 CAAGAAGCTTCTCCAGAAATCAACAGACCTAGACATAGCAAAAGCAATTTATGCA 2441
QY 2404 CTACTCTTCAGTCAGATCA 2422
Db 2442 CCACAAATTCAGAAATGA 2460

```

Search completed: August 18, 1999, 15:12:34
Job time: 2582 sec

OM nucleic - nucleic search, using sw model

11635.497 million cell updates/sec

Title: BUS-09-001-039A-46
 Perfect score: 4832
 Sequence: 1 CTCGAGCTAAGATATTTA.....ATTATGATGTTGCGGCCGC 4832

Scoring table: IDENTITY_NUC

Searched: 311585 seqs, 125096042 residues

Database : N_Geneseq_36:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4832	100.0	4832	1	V15338	Human Factor VIII
2	4832	100.0	4832	1	V19581	Human factor VIII
3	4470.8	94.6	9354	1	T73164	CDNA encoding huma
4	4474.2	92.6	4629	1	Q76016	B-domain deleted F
5	4462	92.3	4670	1	V23339	Human Factor-VIII
6	4322.6	89.5	4830	1	N8154	Human Factor-VIII
7	4216.6	87.3	4275	1	N90654	Human Factor VIII
8	4215.2	87.2	4272	1	N80446	DNA encoding /740 A
9	4213.6	84.2	4545	1	N80447	Modified factor VII
10	4069.6	80.2	4616	1	N81545	Modified factor VII
11	3877.2	76.5	5035	1	T69811	Human Factor VIII
12	3656.4	66.9	4334	1	V12142	Factor VIII-dB695
13	3233.4	62.9	4451	1	V23288	Human Factor VIII
14	2531.2	52.4	9009	1	Q50189	Synthetic human Fac
15	2468.8	51.1	8967	1	T31031	Human Factor VIII
16	2468.8	51.1	9009	1	V12112	Factor-VIII full-l
17	2468.8	51.1	9009	1	V12112	Homo sapiens facto
18	2468.8	51.1	9068	1	V15359	Human Factor VIII
19	2468.8	51.1	9009	1	V25810	Human Factor VIII
20	2468.8	51.1	9080	1	V19580	Human factor III e
21	2467.2	51.1	8241	1	N81439	Human factor III e
22	2467.2	51.1	8975	1	N81096	Factor VIII cDNA 1
23	2467.2	51.1	8975	1	N50054	CDNA sequence encc
24	2464	51.0	9009	1	T61548	Human factor VIII
25	2464	51.0	9009	1	V18884	Factor VIII:C (Argy
26	2348.6	48.6	7440	1	N60689	Homo sapiens facto
27	2302.6	47.7	7059	1	N81543	Sequence encoding
28	2296.4	47.5	6300	1	T03571	Human Factor VIII
29	2295	47.5	7053	1	T51357	Factor-VIII cDNA.
30	2292.6	47.4	7440	1	N50375	Factor VIII:C codi
31	2290.2	47.4	7056	1	Q66615	DNA sequence encco
32	1812	37.5	7032	1	V99801	Sequence of human
33	1806.4	37.4	7493	1	V12115	Canine factor VIII
34	1806.4	37.4	7493	1	V25812	Mus musculus facto
35	1730.2	35.8	6402	1	V12113	Murine factor VIII
36	1240.4	25.7	1623	1	Q66616	Sus scrofa Factor
37	997.8	20.6	3852	1	N50094	Sequence of porcine
38	879.6	18.2	1130	1	Q50187	DNA encoding facto
39	879.6	18.2	1130	1	V12114	Porcine Factor VII
40	879.6	18.2	1130	1	V25811	Sus scrofa factor
41	675.4	14.0	955	1	Q05844	Porcine factor VII
42	675.4	14.0	955	1	N50053	Factor VIII insert
43	480	9.9	480	1	V40051	Sequence of PESVDK
						C2 CDNA SEQ ID NO.

	44	390	8.1	390	1	V40054	D1 cDNA SEQ ID NO:
	45	301.8	6.2	6925	1	T73948	Human Factor V mut

ALIGNMENTS

RESULT	1						
V15338							
ID	V15338	standard; DNA;	4832 BP.				
AC	V15338;						
DT	20-JUL-1998	(first entry)					
DE	Human Factor VIII SQN deletion mutant DNA.						
KW	Factor VIII; blood clotting; haemophilia A; gene therapy;						
OS	Homo sapiens.						
OS	Synthetic.						
FT	Key	Location/Qualifiers					
FT	CDS	72..4445					
FT	CDS	/tag= a					
PN	M09800542-A2.						
PD	08-JAN-1998.						
PR	02-JUN-1997;	U11785					
PR	04-JUN-1997;	US-869309.					
PR	03-JUL-1996;	US-645601.					
PR	13-AUG-1996;	US-696381.					
PA	(CHIR) CHIRON CORP.						
PI	Allen JR, Barber JR, Boder M, Chang SM, Chong K,						
PI	De la Vega D, Depolo NJ, Greengard J, Hsu DC, Ibanez CE,						
PI	Jolly DJ, Mittelstaedt DM, Prussak CE, Respess JG;						
DR	WP1: 98-086967/08.						
P-PSDB:	W4372.						
PT	New replication defective recombinant retroviruses - which express B						
PT	domain-deleted human factor VIII or human factor IX for the						
PT	treatment of haemophilia						
PS	Claim 6: Page 174-175; 236pp; English.						
CC	This DNA sequence includes a coding region for the B domain						
CC	deletion mutant SQN (see W4372) of human Factor VIII. The SQN						
CC	mutant is created by fusing Ser-743 to Gln-1638 of native Factor						
CC	VIII (see W4373) to form a Ser-Gln-Asn (SQN) link between the A2						
CC	and A3 Factor VIII domains. When compared to plasmid-derived						
CC	Factor VIII, the SQN deletion does not influence the in vivo						
CC	pharmacokinetics, but the reduced size of the molecule appears to						
CC	decrease proteolytic degradation. The invention relates to						
CC	preparations of replication defective recombinant retrovirus (RV)						
CC	expressing a B domain-deleted human Factor VIII protein, where the						
CC	recombinant RV is capable of infecting human cells, is resistant						
CC	to degradation by human complement and is capable of inducing						
CC	long-term (at least 30 days and up to 6 months or longer						
CC	post-injection) systemic expression of Factor VIII when						
CC	administered to a haemophilia A patient.						
SC	Sequence 4832 BP;	1381 A;	1089 C;	1055 G;	1307 T;		

Query Match	100.0%;	Score 4832;	DB 1;	Length 4832;			
Best Local Similarity	100.0%;	Pred. NO. 0;					
Matches 4832;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;			

OY	1	CTCGAGCTAAAGATATTATTAGAGAAGATAATTAACCTTTTGCTTCTCAGTTGAACATTTGT	60
Db	1	CTCGAGCTAAAGATATTATTAGAGAAGATAATTAACCTTTTGCTTCTCAGTTGAACATTTGT	60
OY	61	AGCAATTAAGTCATGCACAATTAGAGCTCTCCACCCTGCTTCTTTCTGTGGCTTTTGCATTCT	120
Db	61	AGCAATTAAGTCATGCACAATTAGAGCTCTCCACCCTGCTTCTTTCTGTGGCTTTTGCATTCT	120
OY	121	GCTTAGAGGCCCGAAGAAGATACTACTCGTGGGGGCGAGTGGAACTGCATAGGGGACTATTAGC	180
Db	121	GCTTAGAGGCCCGAAGAAGATACTACTCGTGGGGGCGAGTGGAACTGCATAGGGGACTATTAGC	180
OY	181	AAAGTATCTGGGTAGAGCTGCTGTGAGCAGCAAGATTCTCTTAGAGTGGCCAAAATCTT	240
Db	181	AAAGTATCTGGGTAGAGCTGCTGTGAGCAGCAAGATTCTCTTAGAGTGGCCAAAATCTT	240

QY 241 TTCCATTCAACACCTGCTGCTGTACAAAAAGACTCTTTGTAGATTCAGGATCACC 300
| | | | |
Db 241 TTCCATTCAACACCTGCTGCTGTACAAAAAGACTCTTTGTAGATTCAGGATCACC 300
QY 301 TTTTCAACATCGCTAGCAGGACCCCTGATGGGTCTGTAGTCTTACCATCCAGG 360
| | | | |
Db 301 TTTTCAACATCGCTAGCAGGACCCCTGATGGGTCTGTAGTCTTACCATCCAGG 360
QY 361 CTGAGGTTTATGATACAGTGTCTATTACACTTAAGAAACATGGCTTCCATCTGTGATC 420
| | | | |
Db 361 CTGAGGTTTATGATACAGTGTCTATTACACTTAAGAAACATGGCTTCCATCTGTGATC 420
QY 421 TTCACTGCTGTGGTGTATCTTACAGGAACTCTGAGGAGCGTATGATGATAGA 480
| | | | |
Db 421 TTCACTGCTGTGGTGTATCTTACAGGAACTCTGAGGAGCGTATGATGATAGA 480
QY 481 CCAGTCAAAAGGAGAAAGATGATTAAGTCTTCCCTGTGAGGACCATACATATGCT 540
| | | | |
Db 481 CCAGTCAAAAGGAGAAAGATGATTAAGTCTTCCCTGTGAGGACCATACATATGCT 540
QY 541 GGGAGGTCCTGGAAGAGATGCTCAATGGCTCTGACCCAGTGTGCTTACTACTAT 600
| | | | |
Db 541 GGGAGGTCCTGGAAGAGATGCTCAATGGCTCTGACCCAGTGTGCTTACTACTAT 600
QY 601 ATCTTCTCATGTGACCTGTATAAAGACTTGAATTCAGGCTCATGAGCCCTACTAG 660
| | | | |
Db 601 ATCTTCTCATGTGACCTGTATAAAGACTTGAATTCAGGCTCATGAGCCCTACTAG 660
QY 661 TATGTAGAGAGAGGAGTGTGGCCAAAGAAACACAGACCTTGCAAAATTTATACTAC 720
| | | | |
Db 661 TATGTAGAGAGAGGAGTGTGGCCAAAGAAACACAGACCTTGCAAAATTTATACTAC 720
QY 721 TTTTGTGTATTTTATGAGAGGAAAAAGTTGGCACTAGAAACAAAGACTCCTTGATGC 780
| | | | |
Db 721 TTTTGTGTATTTTATGAGAGGAAAAAGTTGGCACTAGAAACAAAGACTCCTTGATGC 780
QY 781 AGGATAGGAGTGTGCTGCTGCTGCGGCTTAAATGCAACAGTCAATGGTTATG 840
| | | | |
Db 781 AGGATAGGAGTGTGCTGCTGCTGCGGCTTAAATGCAACAGTCAATGGTTATG 840
QY 841 TAAACAGGTCCTGCCAGGTCATGATGATGACACAGAAATAGTCTATTTGGCATGTGA 900
| | | | |
Db 841 TAAACAGGTCCTGCCAGGTCATGATGATGACACAGAAATAGTCTATTTGGCATGTGA 900
QY 901 TTGGAATGGGACACACTCCTGAGTGTACATATTCCTCGAAGTGCACACATTTCTTG 960
| | | | |
Db 901 TTGGAATGGGACACACTCCTGAGTGTACATATTCCTCGAAGTGCACACATTTCTTG 960
QY 961 TGAGGACCATGCGCAGGCGCTCTGGAATCTCGCCAAATACCTTCTTACTGCTCAA 1020
| | | | |
Db 961 TGAGGACCATGCGCAGGCGCTCTGGAATCTCGCCAAATACCTTCTTACTGCTCAA 1020
QY 1021 CACTCTTGATGACCTTGGACACTTTCCTGTTTGTCTATTCCTTCCACCAACATG 1080
| | | | |
Db 1021 CACTCTTGATGACCTTGGACACTTTCCTGTTTGTCTATTCCTTCCACCAACATG 1080
QY 1081 ATGGCATGGAAGTATGTCAAAGTAGACAGTGTCCAGAGGAAACCCCACTACGATGA 1140
| | | | |
Db 1081 ATGGCATGGAAGTATGTCAAAGTAGACAGTGTCCAGAGGAAACCCCACTACGATGA 1140
QY 1141 AAAATATGAGAGAGGAGAGACTATGATGATGATCTTACTGATTCGAAATGGATGTG 1200
| | | | |
Db 1141 AAAATATGAGAGAGGAGAGACTATGATGATGATCTTACTGATTCGAAATGGATGTG 1200
QY 1201 TAGAGTTTATGATGACAACTCTCTTCTTATTCCAAATTCGCTCAGTTGCCAAGAGC 1260
| | | | |
Db 1201 TAGAGTTTATGATGACAACTCTCTTCTTATTCCAAATTCGCTCAGTTGCCAAGAGC 1260
QY 1261 ATCTTAATAAATTGGGTACATTTACATTTGCTGCTGAAGAGGAGGACTGAGTCTCCCT 1320
| | | | |
Db 1261 ATCTTAATAAATTGGGTACATTTACATTTGCTGCTGAAGAGGAGGACTGAGTCTCCCT 1320
QY 1321 TAGTCTCCGCCCCGATGACAGAGATTATTAATAAGTCAATATTTGAACAATGGCCTCAGC 1380
| | | | |
Db 1321 TAGTCTCCGCCCCGATGACAGAGATTATTAATAAGTCAATATTTGAACAATGGCCTCAGC 1380
QY 1381 GGAATGGTGAAGAGTACAAAAAAGTCCGATTTATGGCATACACAGATGAAACCTTTAAG 1440
| | | | |
Db 1381 GGAATGGTGAAGAGTACAAAAAAGTCCGATTTATGGCATACACAGATGAAACCTTTAAG 1440
QY 1441 CTCGTAAAGCTATTCAGCATGATCAGGAATCTTGGGACCTTACTTATGAGGGAAGTTG 1500
| | | | |
Db 1441 CTCGTAAAGCTATTCAGCATGATCAGGAATCTTGGGACCTTACTTATGAGGGAAGTTG 1500
QY 1501 GAGACACTGTTGATTTATTTAAGATCAAGCAAGCAGACCATATTAATCTTACCTTC 1560
| | | | |
Db 1501 GAGACACTGTTGATTTATTTAAGAAATCAAGCAGACCATATTAATCACTACCTTC 1560
QY 1561 ACGAATCACTGATGTCCTGCTTGTATTCAGAGGAGATTAACAAAGGTGTAACAAAT 1620
| | | | |
Db 1561 ACGAATCACTGATGTCCTGCTTGTATTCAGAGGAGATTAACAAAGGTGTAACAAAT 1620
QY 1621 TGAAGATTTTCCAAATTCGCCAGGAGAAATATTCAAATATTAATGAGACGTAGTAG 1680
| | | | |
Db 1621 TGAAGATTTTCCAAATTCGCCAGGAGAAATATTCAAATATTAATGAGACGTAGTAG 1680
QY 1681 AAGATGGGCACTAATACAGATCCTCGGTGCTGACCCGCTATTTACTTACTGCTTA 1740
| | | | |
Db 1681 AAGATGGGCACTAATACAGATCCTCGGTGCTGACCCGCTATTTACTTACTGCTTA 1740
QY 1741 AATAGGAGAGATATCAGCTTACAGACTATGACCTCTCCTCATCTGCTCAAAAGAT 1800
| | | | |
Db 1741 AATAGGAGAGATATCAGCTTACAGACTATGACCTCTCCTCATCTGCTCAAAAGAT 1800
QY 1801 CTGTAGATCAAAAGAGAAACAGATTAATGTACAGAAAGAGAAATGCAATCTGTTTCG 1860
| | | | |
Db 1801 CTGTAGATCAAAAGAGAAACAGATTAATGTACAGAAAGAGAAATGCAATCTGTTTCG 1860
QY 1861 TATTTGATGAGAACCGAGCTGTACCTCAGAGAAATTAACAGCTTCTCCCAATC 1920
| | | | |
Db 1861 TATTTGATGAGAACCGAGCTGTACCTCAGAGAAATTAACAGCTTCTCCCAATC 1920
QY 1921 CAGCTGAGTGCAGCTTGGAGATCCAGAGTTCCAGGCTCCAAATCAGACACATCA 1980
| | | | |
Db 1921 CAGCTGAGTGCAGCTTGGAGATCCAGAGTTCCAGGCTCCAAATCAGACACATCA 1980
QY 1981 ATGCTATTTTTTTATAGATTTGCACTGTGCACTTGTGTCATGAGTGCATCTGCT 2040
| | | | |
Db 1981 ATGCTATTTTTTTATAGATTTGCACTGTGCACTTGTGTCATGAGTGCATCTGCT 2040
QY 2041 ACATTTCTAAGCATTTGAGACAGACTGACTTCTTCTGCTCTTCTGATATTAATCT 2100
| | | | |
Db 2041 ACATTTCTAAGCATTTGAGACAGACTGACTTCTTCTGCTCTTCTGATATTAATCT 2100
QY 2101 TCAAAACAAAAATGGTCTATGAGAGACACATCACCTTATCCATTCAGAGAGAACTG 2160
| | | | |
Db 2101 TCAAAACAAAAATGGTCTATGAGAGACACATCACCTTATCCATTCAGAGAGAACTG 2160
QY 2161 TCTTATGTCGATGGAAGAAACCCAGGTCATGATGATTCGCGGTCACAACTCAGACTTTC 2220
| | | | |
Db 2161 TCTTATGTCGATGGAAGAAACCCAGGTCATGATGATTCGCGGTCACAACTCAGACTTTC 2220
QY 2221 GGAACAGAGGATGAGCGGCTTACTGAAAGTTTCTAGTTGACAGAGAACTGGTAT 2280
| | | | |
Db 2221 GGAACAGAGGATGAGCGGCTTACTGAAAGTTTCTAGTTGACAGAGAACTGGTAT 2280
QY 2281 ATTACAGAGACATTTATGAGATATTTGAGATATTTGCACTACTGATTAACATGCAATG 2340
| | | | |
Db 2281 ATTACAGAGACATTTATGAGATATTTGAGATATTTGCACTACTGATTAACATGCAATG 2340
QY 2341 AACCAAGAGCTTCTCCAGAAACCCAGCTTGTGAAGCCATCAACGGGAATTAATCTC 2400
| | | | |
Db 2341 AACCAAGAGCTTCTCCAGAAACCCAGCTTGTGAAGCCATCAACGGGAATTAATCTC 2400
QY 2401 GTACTACTCTTCAGTCAATCAAGAGAAATTTGACTATGATGATACCATATCAGTTGAAA 2460
| | | | |

Db	2401	GTACTACTCTTCACAGTACAGATCAGAGGAATTTGACTATGATATACATATCATACATTCAGTTGAAA	2460
Qy	2461	TGAAGAGAGAGATTTTGGACATTTATGATGAGATGAAAATCAGAGCCCCCGCAGCTTTC	2520
Db	2461	TGAAGAGAGAGATTTTGGACATTTATGATGAGATGAAAATCAGAGCCCCCGCAGCTTTC	2520
Qy	2521	AAAAGAAAACACGACACTATTTTATTTGCTGCGAGTGGAGAGGCTCTGGGATTTATGGGATGA	2580
Db	2521	AAAAGAAAACACGACACTATTTTATTTATTTGCTGCGACATGGAGAGGCTCTGGGATTTATGGGATGA	2580
Qy	2581	GTACCTCCCAATGTTCTTAAGAAAACAGGGCTCAGAGTGGAGCTGCCCTCAGTTCACAGA	2640
Db	2581	GTACCTCCCAATGTTCTTAAGAAAACAGGGCTCAGAGTGGAGCTGCCCTCAGTTCACAGA	2640
Qy	2641	AAGTTGTTTCCAGAGAA'TTACTGATGCTGCTTACTCAGCCCTTAATACGTTGAGAAC	2700
Db	2641	AAGTTGTTTCCAGAGAA'TTACTGATGCTGCTTACTCAGCCCTTAATACGTTGAGAAC	2700
Qy	2701	TAAATGAAACTTTTGGGACCTCCTGGGGCCATATATAGACACAAAGTTGAAGATATATCA	2760
Db	2701	TAAATGAAACTTTTGGGACCTCCTGGGGCCATATATAGACACAAAGTTGAAGATATATCA	2760
Qy	2761	TGTAATCTTTCAGAAATCAGAGCCTCTCGTCCCATATATAGACACAAAGTTGAAGATATATCA	2820
Db	2761	TGTAATCTTTCAGAAATCAGAGCCTCTCGTCCCATATATAGACACAAAGTTGAAGATATATCA	2820
Qy	2821	ATGAGAGAGATCAGAGGCAAGAGACAGAAACTATGAAAAAATTTGTCAAGCCTAATGAAA	2880
Db	2821	ATGAGAGAGATCAGAGGCAAGAGACAGAAACTATGAAAAAATTTGTCAAGCCTAATGAAA	2880
Qy	2881	CCAAAACTTACTTTTGGAAAGTCAACATCATATGCGACCCCACTAAGATGATTTGACT	2940
Db	2881	CCAAAACTTACTTTTGGAAAGTCAACATCATATGCGACCCCACTAAGATGATTTGACT	2940
Qy	2941	GCAAGAGCTGGGCTTATTTCTGATGATGAGCTGGAAAAAATATGCACTCAGGCCGA	3000
Db	2941	GCAAGAGCTGGGCTTATTTCTGATGATGAGCTGGAAAAAATATGCACTCAGGCCGA	3000
Qy	3001	TTGAGACCCCTTGTGCTGTCCACACTAACACACTGAACCCCTCATGGGAGACAAAGTGA	3060
Db	3001	TTGAGACCCCTTGTGCTGTCCACACTAACACACTGAACCCCTCATGGGAGACAAAGTGA	3060
Qy	3061	CAGTACAGGAA'TTGGCTGTCTTTTACACACTTTTATGAGACCAAAAGCTGTACTTCA	3120
Db	3061	CAGTACAGGAA'TTGGCTGTCTTTTACACACTTTTATGAGACCAAAAGCTGTACTTCA	3120
Qy	3121	CTGAAATATGSAAGAAACTGCAAGGGCTCCCTGCAATATTCAGATGGAAGTCCCACTT	3180
Db	3121	CTGAAATATGSAAGAAACTGCAAGGGCTCCCTGCAATATTCAGATGGAAGTCCCACTT	3180
Qy	3181	TTAAGAGAA'TTATCCGTTCTCATGCAATCATGGCTATATGATGATACTACCTGGCT	3240
Db	3181	TTAAGAGAA'TTATCCGTTCTCATGCAATCATGGCTATATGATGATACTACCTGGCT	3240
Qy	3241	TAGTAATGEGCTCAGAGATCAAGAGATTCGATGATCTGCTCAGCATGGGCAGCAATGAAA	3300
Db	3241	TAGTAATGEGCTCAGAGATCAAGAGATTCGATGATCTGCTCAGCATGGGCAGCAATGAAA	3300
Qy	3301	ACATCCACTTCAATTCATTTAGTGGACATGTGTTCACTGTACGAAAAAAAGGAGGTATA	3360
Db	3301	ACATCCACTTCAATTCATTTAGTGGACATGTGTTCACTGTACGAAAAAAAGGAGGTATA	3360
Qy	3361	AAATGGACATGTAACAATCTATATCAGAGGTGTTTTAGACAGTGGAAATGTTACCATCA	3420
Db	3361	AAATGGACATGTAACAATCTATATCAGAGGTGTTTTAGACAGTGGAAATGTTACCATCA	3420
Qy	3421	AAGCTGGAATTTGGCGGGGTGGAATGCCCTTAATTTGGCAGCATCTACATGCTGGGATGACA	3480
Db	3421	AAGCTGGAATTTGGCGGGGTGGAATGCCCTTAATTTGGCAGCATCTACATGCTGGGATGACA	3480
Qy	3481	CACCTTTTTCGGGTATAGAGCAATATAGTGTCAAGCTCCCTGGGAATGGCTTGTGGACACA	3540
Db	3481	CACCTTTTTCGGGTATAGAGCAATATAGTGTCAAGCTCCCTGGGAATGGCTTGTGGACACA	3540

Db	3481	CACCTTTTTCGTGGTACAGCAATAAAGTGTACAGCTCCCTGGGAATGGCTTCTGTGACACA	3540
Oy	3541	TTAGAGATTTTTCAGATTACAGCTTCAGGACCAATATGACAGTGGGCCCCCAAGCTGGCCA	3600
Db	3541	TTAGAGATTTTTCAGATTACAGCTTCAGGACCAATATGACAGTGGGCCCCCAAGCTGGCCA	3600
Oy	3601	GACTTCATATATCCGAGTCAATATATGCTTGGAGACCAAGAGCCCTTTTCTTGGATCA	3660
Db	3601	GACTTCATATATCCGAGTCAATATATGCTTGGAGACCAAGAGCCCTTTTCTTGGATCA	3660
Oy	3661	AGGTGATCTGTGGGACCAACATGATTATTCACGGCATCAAGACCCAGAGGTGGCCGTGACA	3720
Db	3661	AGGTGATCTGTGGGACCAACATGATTATTCACGGCATCAAGACCCAGAGGTGGCCGTGACA	3720
Oy	3721	AGTTCTCCAGCCTCTACATCTCTCAGTTTATCATCATCTATATAGTCTTGTATGGAGAAGT	3780
Db	3721	AGTTCTCCAGCCTCTACATCTCTCAGTTTATCATCATCTATATAGTCTTGTATGGAGAAGT	3780
Oy	3781	GGCAGACTTATCGAGGAATTCACCTGGAACTTAATGTCCTTTTGGCAATGTGATTT	3840
Db	3781	GGCAGACTTATCGAGGAATTCACCTGGAACTTAATGTCCTTTTGGCAATGTGATTT	3840
Oy	3841	CATCTGGGATTAACACAAATTTTAAACCCCTCAATTTGCTCTGTATACATCCGTTTGC	3900
Db	3841	CATCTGGGATTAACACAAATTTTAAACCCCTCAATTTGCTCTGTATACATCCGTTTGC	3900
Oy	3901	ACCCAACTCATATATAGCATTCGAGCACTCTTCGATGAGTGTATGGGCTGTGATTTAA	3960
Db	3901	ACCCAACTCATATATAGCATTCGAGCACTCTTCGATGAGTGTATGGGCTGTGATTTAA	3960
Oy	3961	ATACTTGCAGATGCCATTGGGAATGGAAGTAAAGCAATATACAGATGCACAGATTACTG	4020
Db	3961	ATACTTGCAGATGCCATTGGGAATGGAAGTAAAGCAATATACAGATGCACAGATTACTG	4020
Oy	4021	CTTATCTCTACTTACCAATATGTTGGCAACCGGTGCTCTCAAAAGCTGCACCTTCAAC	4080
Db	4021	CTTATCTCTACTTACCAATATGTTGGCAACCGGTGCTCTCTCAAAAGCTGCACCTTCAAC	4080
Oy	4081	TCCAAGGGAGAGTAAATGCTTGAGACCTCAGTGTATATCCAAAGAGTGGCTGCACAG	4140
Db	4081	TCCAAGGGAGAGTAAATGCTTGAGACCTCAGTGTATATCCAAAGAGTGGCTGCACAG	4140
Oy	4141	TGCACTTCCAGAGACATGAAAGTACAGAGGTAACTACTCAGGAGTAAATATCTCTGC	4200
Db	4141	TGCACTTCCAGAGACATGAAAGTACAGAGGTAACTACTCAGGAGTAAATATCTCTGC	4200
Oy	4201	TTACACAGATGTATGTGAAGAGTTCCTCATCTCCAGCAGTCAAAATGGCATTAGTGTGA	4260
Db	4201	TTACACAGATGTATGTGAAGAGTTCCTCATCTCCAGCAGTCAAAATGGCATTAGTGTGA	4260
Oy	4261	CTCTCTTTTTTCAGAAATGGCAAAATGAAGTTTTTCAGGGAATCAAGCTCCTTCAAC	4320
Db	4261	CTCTCTTTTTTCAGAAATGGCAAAATGAAGTTTTTCAGGGAATCAAGCTCCTTCAAC	4320
Oy	4321	CTGTGTGAACCTCTAGACCCACCGTTACTGTACTGTACTTGAATTCACCCCCAGA	4380
Db	4321	CTGTGTGAACCTCTAGACCCACCGTTACTGTACTGTACTTGAATTCACCCCCAGA	4380
Oy	4381	GTTGGGTGCACAGATGTCCTTAGATGAGAGTTTGGGGTGCAGAGCACAGACCTCT	4440
Db	4381	GTTGGGTGCACAGATGTCCTTAGATGAGAGTTTGGGGTGCAGAGCACAGACCTCT	4440
Oy	4441	ACTAGAGGTGGCCACTGCAGACCTGCACCTGCGCTCACTCTCCCTCTCAAGCTCCAG	4500
Db	4441	ACTAGAGGTGGCCACTGCAGACCTGCACCTGCGCTCACTCTCCCTCTCAAGCTCCAG	4500
Oy	4501	GCACTGTCCCTCCCTGGCTTGCCTTCTACCTTGTGTCTAAATCTTAGCAGACACTGCTT	4560
Db	4501	GCACTGTCCCTCCCTGGCTTGCCTTCTACCTTGTGTCTAAATCTTAGCAGACACTGCTT	4560
Oy	4561	GAACCTCTCTGAATTAATCATCTACAGCTCTGTGATTTCTTGGTGGGGGCCAGAGAGGT	4620
Db	4561	GAACCTCTCTGAATTAATCATCTACAGCTCTGTGATTTCTTGGTGGGGGCCAGAGAGGT	4620

OY	4621	GCATCCATTAACTACTACTCTTCCACTTAATTTTCGACAGCTGCTCCACAGATTA	CTCTTCC	4680
OY	4621	GCATCCATTAACTACTACTCTTCCACTTAATTTTCGACAGCTGCTCCACAGATTA	CTCTTCC	4680
OY	4681	TTCCAAATATTAACCTAGGCAAAAAGAGAGAGAGAAACCTGCATGAAGAATCTTCC	CTG	4740
Db	4681	TTCCAAATATTAACCTAGGCAAAAAGAGAGAGAGAAACCTGCATGAAGAATCTTCC	CTG	4740
OY	4741	AAATATTGGGCTCTCAGAGTCACCACTTCTCTGTTGTGAAAAAATATGATGAAC	4800	
OY	4741	AAATATTGGGCTCTCAGAGTCACCACTTCTCTGTTGTGAAAAAATATGATGAAC	4800	
OY	4801	TTTGAAAAAGATATTATATGATGTCGGCCGC	4832	
Db	4801	TTTGAAAAAGATATTATATGATGTCGGCCGC	4832	
PS	19581	VI9581 standard: cDNA: 4832 BP.		
DE	06-AUG-1998	(first entry)		
KW	Human factor VIII beta-domain deleted SQN deletion cDNA sequence.			
KW	Replication defective; recombinant retrovirus; RRV, therapeutic protein;			
KW	hepatitis; thrombosis; hypercoagulable disorder; liver disease; human;			
KW	hepatitis; thalassemia; phenylketonuria; Lesch-Nyhan syndrome; diabetes;			
KW	cystic fibrosis; Duchenne's Muscular Dystrophy; hypercholesterolemia;			
KW	hypopituitarism; adenine deaminase deficiency; HIV infection; anaemia;			
KW	Gucher's syndrome; high blood pressure; Alzheimer's disease; autoimmune;			
KW	Inflammatory disease; factor VIII; ss.			
OS	Homo sapiens.			
PN	MO9800541-A2.			
PD	08-JAN-1998.			
PE	02-JUL-1997: U11784.			
PR	04-JUN-1997: US-869309.			
PR	03-JUL-1996: US-645601.			
PR	13-AUG-1996: US-696381.			
PA	(CHIR) CHIRON CORP.			
PI	Allen JR, Barber JR, Boder M, Chang SMW, Chong K,			
PI	De la Vega D, Depolton J, Greenard J, Hsu DC, Ihaner CF,			
PI	Jolly DJ, Lee R, Mittelstaedt DM, Prussak CE, Respass JG;			
DR	WPI: 98-086966/08.			
DR	P-PSDB: W46246.			
PT	New replication defective recombinant retro-viruses - which can be			
PT	administered to provide long term systemic expression of therapeutic			
PT	protein in blood, useful in, e.g. treating hyper-coagulable			
PT	disorders			
PS	Example 28: Pages 210-213; 272pp; English.			
PS	This cDNA encodes the beta-domain deleted SQN deletion protein of human			
CC	factor VIII. This is used in the construction of recombinant retroviral			
CC	vectors expressing human factor VIII. The invention provides the			
CC	preparation of replication defective recombinant retrovirus (RRV)			
CC	expressing a therapeutic protein. The RRV preparation is resistant to			
CC	degradation by human complement and is capable of inducing long term			
CC	systemic expression of the therapeutic protein when administered			
CC	intravenously to a human. The long term systemic expression results in a			
CC	measurable level of the therapeutic protein being produced in the blood			
CC	of the human for a period of at least 30 days after the administration of			
CC	the RRV vector preparation. RRV's can be used for in vivo delivery of			
CC	therapeutic protein to treat, e.g. haemophilia A, haemophilia B,			
CC	thrombosis, hypercoagulable disorders, liver diseases such as hepatitis,			
CC	disorders such as thalassemia, phenylketonuria, Lesch-Nyhan syndrome,			
CC	severe combined immunodeficiency (SCID), cystic fibrosis, Duchenne's			
CC	Muscular Dystrophy, inherited emphysema, familial hypercholesterolemia,			
CC	diabetes, hypopituitarism, adenine deaminase deficiency, alpha1-			
CC	antitrypsin deficiency, Gucher's syndrome, anaemia, infections such as			
CC	HIV infection, high blood pressure, Alzheimer's disease, autoimmune or			
CC	inflammatory disease or graft versus host disease. RRV's are capable of			
CC	surviving inactivation in human serum thereby allowing efficient gene			
CC	transfer over prolonged periods of time.			
CC	Sequence 4832 BP: 1381 A; 1089 C; 1055 G; 1307 T;			
CC	Sequence 4832 BP: 1381 A; 1089 C; 1055 G; 1307 T;			

Query Match	100.0%	Score 4832.	DB 1.	Length 4832.	
Best Local Similarity	100.0%	Pred. No. 0.			
Matches 4832.	Conservative	0.	Mismatches	0.	Indels
					Gaps
0.					
QY	1	CTCGAGCTTAAATATTTTATGAGAGAAATTAACCTTTTCTCTCCAGTGAACATTTGT	60		
Db	1	CTCGAGCTTAAATATTTTATGAGAGAAATTAACCTTTTCTCTCCAGTGAACATTTGT	60		
QY	61	AGCATTAATCATGCATAATAGAGCTCTCCACTGCTTCTTCTGTGCGCTTTTGGCATTTCT	120		
Db	61	AGCATTAATCATGCATAATAGAGCTCTCCACTGCTTCTTCTGTGCGCTTTTGGCATTTCT	120		
QY	121	GCTTTAGTCCCAAGAAAGATCTACTGCTGGGGGACAGTGGAAATGTCATGGGACTTATATGC	180		
Db	121	GCTTTAGTCCCAAGAAAGATCTACTGCTGGGGGACAGTGGAAATGTCATGGGACTTATATGC	180		
QY	181	AAAGTATCTCGGTGAGCTGCTGTGGAAGCAATTTCTCTAGAGTCCCAATATCTT	240		
Db	181	AAAGTATCTCGGTGAGCTGCTGTGGAAGCAATTTCTCTAGAGTCCCAATATCTT	240		
QY	241	TTTCATTTCAACACCTTCATGCTGCTGTCACAAAAGACTCTGTTGTAGAAATTCACGGATCAC	300		
Db	241	TTTCATTTCAACACCTTCATGCTGCTGTCACAAAAGACTCTGTTGTAGAAATTCACGGATCAC	300		
QY	301	TTTTAACATCCCTTAAGCCAAAGCCACCCTGGATGGGTGCTAGTGGTCTTACATTCACG	360		
Db	301	TTTTAACATCCCTTAAGCCAAAGCCACCCTGGATGGGTGCTAGTGGTCTTACATTCACG	360		
QY	361	CTGAGGTTTATGATACAGTGGTTCATTACATTAAGAACATGGCTCCGATCCTGTCAATC	420		
Db	361	CTGAGGTTTATGATACAGTGGTTCATTACATTAAGAACATGGCTCCGATCCTGTCAATC	420		
QY	421	TTTCATGCTGTTGGTATATCTTACCTGGAAGGCTCTGAGAGGACTGAATATATGATGATCGA	480		
Db	421	TTTCATGCTGTTGGTATATCTTACCTGGAAGGCTCTGAGAGGACTGAATATATGATGATCGA	480		
QY	481	CCAGTCAAAAGGAGAAAGATGATTAAGTCTTCCCTGGTGGAGGACCATATATATGCT	540		
Db	481	CCAGTCAAAAGGAGAAAGATGATTAAGTCTTCCCTGGTGGAGGACCATATATATGCT	540		
QY	541	GGCAGGCTCTGAAGAAAGATGGTCCATATGGGCTGTGACCCACTGTGGCTTACCTATCTAT	600		
Db	541	GGCAGGCTCTGAAGAAAGATGGTCCATATGGGCTGTGACCCACTGTGGCTTACCTATCTAT	600		
QY	601	ATCTTCTCATGTGACCTGGTAAAGAGCTTGAATTCAGGCTCATTTGGAGCCCTACTAG	660		
Db	601	ATCTTCTCATGTGACCTGGTAAAGAGCTTGAATTCAGGCTCATTTGGAGCCCTACTAG	660		
QY	661	TATGTAGAGAAAGGAGTGTGGCCAAAGAAAGACACAGACCTTGACAAATTTTACTAC	720		
Db	661	TATGTAGAGAAAGGAGTGTGGCCAAAGAAAGACACACACCTTGACAAATTTTACTAC	720		
QY	721	TTTTGTCGTATTTGATTAAGGAGAAATTTGGCACTCGAAGCAAAAGAACTCTGATATGC	780		
Db	721	TTTTGTCGTATTTGATTAAGGAGAAATTTGGCACTCGAAGCAAAAGAACTCTGATATGC	780		
QY	781	AGGATAGGGATGTGCATCTGCTGGGCTGGCTTAAATGACACAGTCAATGGTATATG	840		
Db	781	AGGATAGGGATGTGCATCTGCTGGGCTGGCTTAAATGACACAGTCAATGGTATATG	840		
QY	841	TAAACAGGTCCTCGCAGGTCGTGATTTGGATGCCACAGAAATTCAGTATTTGGCATGTGA	900		
Db	841	TAAACAGGTCCTCGCAGGTCGTGATTTGGATGCCACAGAAATTCAGTATTTGGCATGTGA	900		
QY	901	TTTGGATTTGGACACATCTCTGAATGCACTCAATATTTCTGGAAGGTACACATTTTCTG	960		
Db	901	TTTGGATTTGGACACATCTCTGAATGCACTCAATATTTCTGGAAGGTACACATTTTCTG	960		
QY	961	TGAGGAACCATTCGCGAGGCTCTTGGAAATCTGCCAATATCTTCTTACTGCTCAAA	1020		
Db	961	TGAGGAACCATTCGCGAGGCTCTTGGAAATCTGCCAATATCTTCTTACTGCTCAAA	1020		

QY	1021	CACCTTGATGGACCTTGAGACAGTTCTACGTGTTTGTCATATCTCTTCCACCAACATG	1080
Db	1021	CACCTTGATGGACCTTGAGACAGTTCTACGTGTTTGTCATATCTCTTCCACCAACATG	1080
QY	1081	ATGGCATGGAAGCTTATGCTAAAGTGAAGAGCTGTCCAGAGACCCCACTACGAATGA	1140
Db	1081	ATGGCATGGAAGCTTATGCTAAAGTGAAGAGCTGTCCAGAGACCCCACTACGAATGA	1140
QY	1141	AAATTAATGAAGAAGCGGAAGACTATGATGATGATCTCTACTATCTGCAAAATGATGTGG	1200
Db	1141	AAATTAATGAAGAAGCGGAAGACTATGATGATGATGATCTCTACTATCTGCAAAATGATGTGG	1200
QY	1201	TCAGGTTTGATGATGACACTCTCTCTCTTATATCCAAATTCGCTACAGTGGCAAGAAGC	1260
Db	1201	TCAGGTTTGATGATGACACTCTCTCTCTTATATCCAAATTCGCTACAGTGGCAAGAAGC	1260
QY	1261	ATCCCTAAACCTGGGTGCTACCTACCTATGCTGCTGAAGAGAGACGTGGGACTATGCTCCT	1320
Db	1261	ATCCCTAAACCTGGGTGCTACCTACCTATGCTGCTGAAGAGAGACGTGGGACTATGCTCCT	1320
QY	1321	TAGTCTCTGCCCCCGATGACAGAAAGTTATAAAAGTCAATNTTGAACATGGCCCTCAGC	1380
Db	1321	TAGTCTCTGCCCCCGATGACAGAAAGTTATAAAAGTCAATNTTGAACATGGCCCTCAGC	1380
QY	1381	GGATTGGTAGAGATACAAAAAGTCCGATTTATGCTATACACAGATGAACCTTTAAGA	1440
Db	1381	GGATTGGTAGAGATACAAAAAGTCCGATTTATGCTATACACAGATGAACCTTTAAGA	1440
QY	1441	CTCGGAAGCTATTCACACTGAAATCAAGAAATCTGGGACCTTACTTTATGGGAAGTGG	1500
Db	1441	CTCGGAAGCTATTCACACTGAAATCAAGAAATCTGGGACCTTACTTTATGGGAAGTGG	1500
QY	1501	GAGACACACTGTTGATATATTTTAAGATACAGCAAGCAGACATATAACTCTACCTC	1560
Db	1501	GAGACACACTGTTGATATATTTTAAGATACAGCAAGCAGACATATAACTCTACCTC	1560
QY	1561	ACGGATCTACTGATGTCCTCTCTTTGTTATCAAGAGATTACCAAAAGGTGTAACACTT	1620
Db	1561	ACGGATCTACTGATGTCCTCTCTTTGTTATCAAGAGATTACCAAAAGGTGTAACACTT	1620
QY	1621	TGAAGGATTTCCAAATCTGCGCAAGGAATATTCAAATATATAAGACAGACTGTAG	1680
Db	1621	TGAAGGATTTCCAAATCTGCGCGAAGGAATATTCAAATATATAAGACAGACTGTAG	1680
QY	1681	AAGATGGGCCAACTAAATCAAGATCTCTCGGTGCTGACCCGCTATTACTGTAGTTCCGTTA	1740
Db	1681	AAGATGGGCCAACTAAATCAAGATCTCTCGGTGCTGACCCGCTATTACTGTAGTTCCGTTA	1740
QY	1741	ATATGAGAGAGATCTAGCTTCAAGACTCATTTGGCCCTCTCTCTCACTGTGCTCAAAAGAT	1800
Db	1741	ATATGAGAGAGATCTAGCTTCAAGACTCATTTGGCCCTCTCTCTCACTGTGCTCAAAAGAT	1800
QY	1801	CTGTAGATCAAAAGGAAACCAATATATGTCAGACAAAGAGATGTCATCCTGTTTCTG	1860
Db	1801	CTGTAGATCAAAAGGAAACCAATATATGTCAGACAAAGAGATGTCATCCTGTTTCTG	1860
QY	1861	TATTTGATGAGAACCGAAGCTGTGACTCTCAGAGATATTAAGAGCTTTCGCCCAATC	1920
Db	1861	TATTTGATGAGAACCGAAGCTGTGACTCTCAGAGATATTAAGAGCTTTCGCCCAATC	1920
QY	1921	CAGCTGAGTGACAGCTTGAGAGATCCAGAGTTCCAAGCCTCAACATCATGCAAGACATCA	1980
Db	1921	CAGCTGAGTGACAGCTTGAGAGATCCAGAGTTCCAAGCCTCAACATCATGCAAGACATCA	1980
QY	1981	ATGCGTATGTTTGGATAGTTTGCAAGTGTGCACTTTGTTGGATGAAGGTGGCATCTGTG	2040
Db	1981	ATGCGTATGTTTGGATAGTTTGCAAGTGTGCACTTTGTTGGATGAAGGTGGCATCTGTG	2040
QY	2041	ACATCTTAAGCATTTGGAGACAGACTACTCTCTTCTGCTCTCTCTCTCTCTCTCTCTCTCT	2100
Db	2041	ACATCTTAAGCATTTGGAGACAGACTACTCTCTTCTGCTCTCTCTCTCTCTCTCTCTCTCT	2100
QY	2101	TCAAAACAAATGCTCTATGAAGACACTCACTCACTCACTCACTCACTCACTCACTCACTCACTG	2160

Db	2101	TCAAACACAAAATGGTCTATGAAGACACACTCACCTATTCCTCCATTCTCAGAGAACTG	2160
QY	2161	TCCTTCATGTCGATGAGAAACCCAGGTCCTATGAGATTCCTGGGGGTGGCCACACACAGACTTC	2220
Db	2161	TCCTTCATGTCGATGAGAAACCCAGGTCCTATGAGATTCCTGGGGGTGGCCACACACAGACTTC	2220
QY	2221	GGAACAGAGGCATGACCGCCTTACTGAAAGTTTCTAGTTGTGTGACAAAGACACTGGTGATT	2280
Db	2221	GGAACAGAGGCATGACCGCCTTACTGAAAGTTTCTAGTTGTGTGACAAAGACACTGGTGATT	2280
QY	2281	ATTACAGGAGACGTTATGAAGATTTTTCACCACTCTTGGTGAATTAACATGCCATTG	2340
Db	2281	ATTACAGGAGACGTTATGAAGATTTTTCACCACTCTTGGTGAATTAACATGCCATTG	2340
QY	2341	AACCAAGAAAGCTTCTCCAGAACCCACAGTCTTGAAGCCCATCAAGGGAAATPACTC	2400
Db	2341	AACCAAGAAAGCTTCTCCAGAACCCACAGTCTTGAAGCCCATCAAGGGAAATPACTC	2400
QY	2401	GTACTACTCTTCAGTCAGATCAAGAGAAATTGACTATGATGTATACATATCAGTTGAA	2460
Db	2401	GTACTACTCTTCAGTCAGATCAAGAGAAATTGACTATGATGTATACATATCAGTTGAA	2460
QY	2461	TGAAGAGAGAAAGATTTTGACATTTATGATGAGATGAAATCAGAGCCCCCGACGTTTC	2520
Db	2461	TGAAGAGAGAAAGATTTTGACATTTATGATGAGATGAAATCAGAGCCCCCGACGTTTC	2520
QY	2521	AAAAGAAAACAGCACTATTTTATGCTGCAATGGAGAGGCTGTGGAAATTAATGGGATGA	2580
Db	2521	AAAAGAAAACAGCACTATTTTATGCTGCAATGGAGAGGCTGTGGAAATTAATGGGATGA	2580
QY	2581	GTAGCTCCCACTATGTTCTAAGAAACAGGGCTCAGAGTGCGACTGTCCCTCAGTTCAAGA	2640
Db	2581	GTAGCTCCCACTATGTTCTAAGAAACAGGGCTCAGAGTGCGACTGTCCCTCAGTTCAAGA	2640
QY	2641	AAAGTTGTTTCCAGGAATTTACTGATGCTCTTACTAGCCCTTATACCGTGGAGAAC	2700
Db	2641	AAAGTTGTTTCCAGGAATTTACTGATGCTCTTACTAGCCCTTATACCGTGGAGAAC	2700
QY	2701	TAAATGAACATTTGGGAGCTCTCTGGGGCCATATATGAGAGAGAAAGTTGAAGATTAATGA	2760
Db	2701	TAAATGAACATTTGGGAGCTCTCTGGGGCCATATATGAGAGAGAAAGTTGAAGATTAATGA	2760
QY	2761	TGGTAACTTTCAGAAATCAGGCGCTCTGTCCTCATTCCTTATCTTAGCCTTATTTCT	2820
Db	2761	TGGTAACTTTCAGAAATCAGGCGCTCTGTCCTCATTCCTTATCTTAGCCTTATTTCT	2820
QY	2821	ATGAGGAAGATCAAGGCGAGGAGCAGACCTTGAAAAAACTTTGTCAAGCCTAATGAAA	2880
Db	2821	ATGAGGAAGATCAAGGCGAGGAGCAGACCTTGAAAAAACTTTGTCAAGCCTAATGAAA	2880
QY	2881	CCAAAACCTACTTTTGGAAAGTGCAAATATATATGACACCCTAAAGATGAGTTGACT	2940
Db	2881	CCAAAACCTACTTTTGGAAAGTGCAAATATATATGACACCCTAAAGATGAGTTGACT	2940
QY	2941	GCAAAAGCTTGCGGCTTATTTCTGATGTGACCTGGAAGAAAGATGSCACTGAGGCTGA	3000
Db	2941	GCAAAAGCTTGCGGCTTATTTCTGATGTGACCTGGAAGAAAGATGSCACTGAGGCTGA	3000
QY	3001	TTGAGACCCCTTGTGTCGCACACTAACAACACTGAAACCTGTCTCATGGGAGACAAGTGA	3060
Db	3001	TTGAGACCCCTTGTGTCGCACACTAACAACACTGAAACCTGTCTCATGGGAGACAAGTGA	3060
QY	3061	CAGTACAGAGAAATTGCTCTGTTTTTTCACACATTTTATATGAGACCAAAACCTGGTACTTA	3120
Db	3061	CAGTACAGAGAAATTGCTCTGTTTTTTCACACATTTTATATGAGACCAAAACCTGGTACTTA	3120
QY	3121	CTGAAATAATGAGAAAGAACTGCAGGGCTCCCTGCATAATCCAGATGGAAGATCCCACTT	3180
Db	3121	CTGAAATAATGAGAAAGAACTGCAGGGCTCCCTGCATAATCCAGATGGAAGATCCCACTT	3180
QY	3181	TTAAAAGAAATTATCGCTTCCTCATGCAATCAATGAGCTATCATATATGATGATACACTACCTGCT	3240
Db	3181	TTAAAAGAAATTATCGCTTCCTCATGCAATCAATGAGCTATCATATATGATGATACACTACCTGCT	3240

[illegible]

FT	misc-feature	4798. .4817	
FT	/tag-1		
FT	/note- '3' near consensus site"		
FT	misc-feature	5023. .5045	
FT	/tag-1		
FT	/note- '3' near consensus site"		
FT	misc-feature	5333. .5355	
FT	/tag- k		
FT	/note- '3' near consensus site"		
FT	misc-feature	5520. .5538	
FT	/tag-1		
FT	/note- '3' near consensus site"		
FT	misc-feature	5604. .5632	
FT	/tag- m		
FT	/note- '3' near consensus site"		
FT	misc-feature	5717. .5745	
FT	/tag- n		
FT	/note- '3' near consensus site"		
FT	misc-feature	6239. .6258	
FT	/tag- o		
FT	/note- '3' near consensus site"		
FT	misc-feature	6658. .6682	
FT	/tag- p		
FT	/note- '3' near consensus site"		
FT	misc-feature	7159. .7176	
FT	/tag- q		
FT	/note- '3' near consensus site"		
FT	misc-feature	7196. .7209	
FT	/tag- r		
FT	/note- '3' near consensus site"		
FT	misc-feature	7289. .7315	
FT	/tag- s		
FT	/note- '3' near consensus site"		
FT	misc-feature	7411. .7429	
FT	/tag- t		
FT	/note- '3' near consensus site"		
FT	misc-feature	7611. .8197	
FT	/tag- u		
FT	/note- "PRE sequence"		
PT	MO9733994-A1.		
PN	18-SEP-1997.		
PD	10-MAR-1997. U03561.		
PE	11-MAR-1996; US-6683839.		
PR	(IMMU-) IMMUNE RESPONSE CORP.		
PA	Bidlingmaier S, Ill CR;		
PI	WPI: 97-470874/43.		
DR	P-PDSB; W23414.		
DT	Vector for increased expression of intronless genes - comprises		
DT	intronless gene with at least one near consensus splice sequence, a		
DT	promoter and at least one viral cis-acting post-transcriptional		
DT	regulatory element		
DT	Example 1; Pages 21-31; 59pp; English.		
DT	The present sequence represents human B-domain deleted factor VIII		
DT	cDNA, and a post-translational regulatory element (PRE) of the		
DT	Hepatitis B virus, which is present 3' of the STOP codon for factor VIII.		
DT	CC Pre sequences have been shown to function in cis to increase the		
DT	steady-state levels of surface gene transcripts by facilitating		
DT	cytoplasmic accumulation of these transcripts. The present sequence		
DT	is part of a novel vector, comprising an intronless gene containing		
DT	1 or more near consensus splice sequences operably linked to a		
DT	promoter sequence so that the gene is transcribed in a cell.		
DT	Intronless gene transcripts which contain near consensus splice site		
DT	sequences are believed to get tied up in the nucleus of the cell where		
DT	splicing occurs, rather than being transported to the cytoplasm where		
DT	they can be translated into proteins. The PRE sequences are transcribed		
DT	along with the gene, causing export of the gene transcript from the		
DT	nucleus into the cytoplasm of the cell. The vector can be used		
DT	to increase the expression of an intronless gene containing at least one		
DT	near consensus splice sites, preferably cDNA encoding a blood coagulation		
DT	factor, particularly Factor VIII or IX. The complex allows the targeted		
DT	delivery of the vector to a specific cell, e.g. hepatocytes when the		
DT	ligand is an asialoglycoprotein which binds the asialoglycoprotein		
DT	receptor present on their surface		

Seq	Sequence	9354 BP:	2506 A:	2239 C:	2161 G:	2448 T:
	Query Match	94.6%;	Score 4570.8;	DB 1;	Length 9354;	
	Best Local Similarity	98.8%;	Pred. No. 0;			
	Matches 4643;	Conservative	0;	Mismatches	12;	Indels 45; Gaps
QY	14 TATTTTGGAGAGAAATTAACTTTTGTCTTCCAGTGTGAACAATTTGTAGCAATAAG---T	70				
DB	2904 TATTTTGGAGAGAAATTAACTTTTGTCTTCCAGTGTGAACAATTTGTAGCAATAAGCCAC	2963				
QY	71 CATCAAAATGAGCGTCGCCACCGTCCTTCTTCTTGCGCCCTTTGGCATTCGTTTAGTGC	130				
DB	2964 CATGTTTATGAGCGTCGCCACCGTCCTTCTTCTTGCGCCCTTTGGCATTCGTTTAGTGC	3023				
QY	131 CACGAGAAGATTAACCTGGGTGCGAGTGGAACTGTCATGAGACATTTATGCAAGAGATCT	190				
DB	3024 CACGAGAAGATTAACCTGGGTGCGAGTGGAACTGTCATGAGACATTTATGCAAGAGATCT	3083				
QY	191 CGGTGAGCGTCCGTGTGGACGCAAGATTTCTCTCTAGAGTCCCAAATCTTTTCCATTCAA	250				
DB	3084 CGGTGAGCGTCCGTGTGGACGCAAGATTTCTCTCTAGAGTCCCAAATCTTTTCCATTCAA	3143				
QY	251 CACCTCAGTGGTGTACAAAAGACTGTGTTTGTAGAAATCAGGATTCACCTTTTCACAT	310				
DB	3144 CACCTCAGTGGTGTACAAAAGACTGTGTTTGTAGAAATCAGGATTCACCTTTTCACAT	3203				
QY	311 CGCTACAGCAAGGCGCACCCGTGATGGGTCTGGAGATCCTTACATCCAGGCGTGAAGTTTA	370				
DB	3204 CGCTACAGCAAGGCGCACCCGTGATGGGTCTGGAGATCCTTACATCCAGGCGTGAAGTTTA	3263				
QY	371 TGATACAGTGGCTATTACACTTAAAGAACATGGCTTCCATCTCTCAGTCTTCATGCTGT	430				
DB	3264 TGATACAGTGGCTATTACACTTAAAGAACATGGCTTCCATCTCTCAGTCTTCATGCTGT	3323				
QY	431 TGGTGTATCTTACTGGAAGACTTCTGAGGAGCTGAATGTGATCAGACCACTTCAAG	490				
DB	3324 TGGTGTATCTTACTGGAAGACTTCTGAGGAGCTGAATGTGATCAGACCACTTCAAG	3383				
QY	491 GGAGAAAGAAAGATATATTAATCTTCCCTGGTGGAAAGCCATACATATATGCTGGCAGGTCT	550				
DB	3384 GGAGAAAGAAAGATATATTAATCTTCCCTGGTGGAAAGCCATACATATATGCTGGCAGGTCT	3443				
QY	551 GAAAGAGAAATGTTCCATATGGCCCTCTGACCCACTGTGCTTACTACTATATCTTTTCA	610				
DB	3444 GAAAGAGAAATGTTCCATATGGCCCTCTGACCCACTGTGCTTACTACTATATCTTTTCA	3503				
QY	611 TGTGGACCTGGTAAAGACACTTGAATTCAGAGCCATATGGAGCCCTACTAGTATGTAGAGA	670				
DB	3504 TGTGGACCTGGTAAAGACACTTGAATTCAGAGCCCTACTAGTATGTAGTATGTAGAGA	3563				
QY	671 AGGAGATCTGGCCCAAGAAAAGACACAGACTTGGACAAATTTATACTATTTTGTCTGT	730				
DB	3564 AGGAGATCTGGCCCAAGAAAAGACACAGACTTGGACAAATTTATACTATTTTGTCTGT	3623				
QY	731 ATTTGATGAAGGGAAGGTTGGGACATCAGAAACAAAGACTCTTGATGCGAGGATAGGA	790				
DB	3624 ATTTGATGAAGGGAAGGTTGGGACATCAGAAACAAAGACTCTTGATGCGAGGATAGGA	3683				
QY	791 TGTGTCATCTGCTCGGGCCCTGGCCCTTAAATTCACACAGATCAATGGTTATGTAAACAGTCT	850				
DB	3684 TGTGTCATCTGCTCGGGCCCTGGCCCTTAAATTCACACAGATCAATGGTTATGTAAACAGTCT	3743				
QY	851 TCTGCGCAGGCTCTGATTTGGATGCGACAGGAATTCAGTCTATGGCATGTGATTTGAATGG	910				
DB	3744 TCTGCGCAGGCTCTGATTTGGATGCGACAGGAATTCAGTCTATGGCATGTGATTTGAATGG	3803				
QY	911 CACACATCTCTGAAGTGCATCATATTTCTCGAAGGTCAACATTTCTGTGAGGAACCA	970				
DB	3804 CACACATCTCTGAAGTGCATCATATTTCTCGAAGGTCAACATTTCTGTGAGGAACCA	3863				
QY	971 TCGCAGGCGTCTTGGAAATCTCGCCAAATTAATTTCTTACTGCTCAACACTCTTGAT	1030				

Dh 3864 TCGCCAGGGCTCTGGAAATCTGCGCAATTAATCTTCTTACTGCTCAACACTCTTGAT 3923
Qy 1031 GGACCTTGGACAGTTCCTCTCTGTTTGTATATCTCTCCACCAACATGATGGCATGA 1090
Dh 3924 GGACCTTGGACAGTTCCTCTCTGTTTGTATATCTCTCCACCAACATGATGGCATGA 3983
Qy 1091 AGCTTATGTCAAAAGTATAGAGCTGTCAGAGAAACCCCACTACGAATGAAAAATTAATGA 1150
Dh 3984 AGCTTATGTCAAAAGTATAGAGCTGTCAGAGAAACCCCACTACGAATGAAAAATTAATGA 4043
Qy 1151 AGAAGCGGAAGCATATGATGATGATCTTACTGATTTGAAATGATGATGATGATGA 1210
Dh 4044 AGAAGCGGAAGCATATGATGATGATCTTACTGATTTGAAATGATGATGATGATGA 4103
Qy 1211 TGATGCAACTCTCTCTCTTATCCAAATGCTGCTAGTTGCCAGAGACATCTTAAAC 1270
Dh 4104 TGATGCAACTCTCTCTCTTATCCAAATGCTGCTAGTTGCCAGAGACATCTTAAAC 4163
1271 TTGGGTACATTTACATTTGCTGTGAAGAGAGAGCTGGACTATGCTCCCTTATGCTCTGC 1330
Dh 4164 TTGGGTACATTTACATTTGCTGTGAAGAGAGAGCTGGACTATGCTCCCTTATGCTCTGC 4223
Qy 1331 CCCCAGTACAGAGATTAATAAAGTCAATATTTGAACAATGGCCCTCAGCGGATTTGAG 1390
Dh 4224 CCCCAGTACAGAGATTAATAAAGTCAATATTTGAACAATGGCCCTCAGCGGATTTGAG 4283
Qy 1391 GAAGTACAAAAAGTCCGATTTATGCAATACAGATGCAACCTTTAGACTCTGTAAGC 1450
Dh 4284 GAAGTACAAAAAGTCCGATTTATGCAATACAGATGCAACCTTTAGACTCTGTAAGC 4343
Qy 1451 TATTCAGCATGATACAGAACTCTGGACCTTACTTTATGGGAAGTTGGAGACACT 1510
Dh 4344 TATTCAGCATGATACAGAACTCTGGACCTTACTTTATGGGAAGTTGGAGACACT 4403
Qy 1511 GTTGATTTATTTAAGATCAAGCAAGACCATTAATACATCTACCTCAGCAATCAC 1570
Dh 4404 GTTGATTTATTTAAGATCAAGCAAGACCATTAATACATCTACCTCAGCAATCAC 4463
Qy 1571 TGATTCGCTCTCTTTGTATTCAGAGAGATTCACAAAGTGTAAACATTTGAAGATTT 1630
Dh 4464 TGATTCGCTCTCTTTGTATTCAGAGAGATTCACAAAGTGTAAACATTTGAAGATTT 4523
Qy 1631 TCCAAATCTGCCAGAGAGAAATATTCAAATATTAATGAGACAGTGTAGAAGTGGCC 1690
Dh 4524 TCCAAATCTGCCAGAGAGAAATATTCAAATATTAATGAGACAGTGTAGAAGTGGCC 4583
1691 AACATAATCAAGATCTCGCTGCTGACCCGCTATTACTCTAGATTTCGTTAATATGAGAG 1750
Dh 4584 AACATAATCAAGATCTCGCTGCTGACCCGCTATTACTCTAGATTTCGTTAATATGAGAG 4643
Qy 1751 AGATCTACCTTCAGAGCATTTGGCCCTCTCTCTCATCTGCTACAAAGAAATCTGATGA 1810
Dh 4644 AGATCTACCTTCAGAGCATTTGGCCCTCTCTCTCATCTGCTACAAAGAAATCTGATGA 4703
Qy 1811 AAGAGAAACAGATTAATGTACAGACAAAGAAATGTCAATCTGTTTCTGATTTGATGA 1870
Dh 4704 AAGAGAAACAGATTAATGTACAGACAAAGAAATGTCAATCTGTTTCTGATTTGATGA 4763
Qy 1871 GAACCGAAGCTGTATCTCAAGAGAAATATACAAAGCTTTTCTCCCAATCCAGCTGAGT 1930
Dh 4764 GAACCGAAGCTGTATCTCAAGAGAAATATACAAAGCTTTTCTCCCAATCCAGCTGAGT 4823
Qy 1931 GCACCTTGAGAGATCCAGATTCACAGCTCTCAACATCATATGACAGCAATGGCTATGT 1990
Dh 4824 GCACCTTGAGAGATCCAGATTCACAGCTCTCAACATCATATGACAGCAATGGCTATGT 4883
Qy 1991 TTTTGATAGTTTGAGATTTGATGATTTGTCATGAGGTGCAATAGTGTATTTAAG 2050
Dh 4884 TTTTGATAGTTTGAGATTTGATGATTTGTCATGAGGTGCAATAGTGTATTTAAG 4933
Qy 2051 CATTTGAGAGACAGACTGACTCTCTTCTGCTTCTCTCTGATATACCTTCAAAACAA 2110
Dh 4944 CATTTGAGAGACAGACTGACTCTCTTCTGCTTCTCTCTGATATACCTTCAAAACAA 5003

Qy 2111 AATGCTATGAGAGACACTACCCCTATTCACATTCCTCAGAGAAACGTCTTATGTC 2170
Dh 5004 AATGCTATGAGAGACACTACCCCTATTCACATTCCTCAGAGAAACGTCTTATGTC 5063
Qy 2171 GATGAAAAACCGGTCTATGATTTCTGGGTGCCAACACTCAGACTTTGGAAACAGAG 2230
Dh 5064 GATGAAAAACCGGTCTATGATTTCTGGGTGCCAACACTCAGACTTTGGAAACAGAG 5123
Qy 2231 CATGACCCCTTACTGAAGGTTTCTAGTTGACAGAAACACTGCTGATTTATTCAGAGA 2290
Dh 5124 CATGACCCCTTACTGAAGGTTTCTAGTTGACAGAAACACTGCTGATTTATTCAGAGA 5183
Qy 2291 CAGTTATGAGATATTTTCAGACTTCTGCTGAGTAAAAACAATGCCATTTAGAACAGAG 2350
Dh 5184 CAGTTATGAGATATTTTCAGACTTCTGCTGAGTAAAAACAATGCCATTTAGAACAGAG 5243
Qy 2351 CTTCCTCCAGAA-----CCGACC 2368
Dh 5244 CTTCCTCCAGAAATTCAGACACACCTAGCAGTACGCAAAAGCAATTTATGCCACCCACC 5303
Qy 2369 AGCTTTGAAAGCCCATCAAGCGGAAATTAACCTGACTACTTTCAGTCAATCAAGAGA 2428
Dh 5304 AGCTTTGAAAGCCCATCAAGCGGAAATTAACCTGACTACTTTCAGTCAATCAAGAGA 5363
Qy 2429 AATTGACTATGATATACCATATCAGTTGAAATGAAAGAGAAATTTGACATTTATGA 2488
Dh 5364 AATTGACTATGATATACCATATCAGTTGAAATGAAAGAGAAATTTGACATTTATGA 5423
Qy 2489 TGAGATGAAATTAAGAGCCCGCCAGCTTTCAAAAGAAACACAGACATTTATTTGC 2548
Dh 5424 TGAGATGAAATTAAGAGCCCGCCAGCTTTCAAAAGAAACACAGACATTTATTTATTC 5483
Qy 2549 TGCAGTGAGAGGCTCTGGGATTTATGGATGATGATGCTCCACATGTTCTTAAGAAACAG 2608
Dh 5484 TGCAGTGAGAGGCTCTGGGATTTATGGATGATGATGCTCCACATGTTCTTAAGAAACAG 5543
Qy 2609 GGCCTCAAGTGGGAGTCTCCCTCAGTTCAAGAAAGTTGTTTCCAGAAATTTACTGATG 2668
Dh 5544 GGCCTCAAGTGGGAGTCTCCCTCAGTTCAAGAAAGTTGTTTCCAGAAATTTACTGATG 5603
Qy 2669 CTCTTTTACAGCCCTTATACGCTGAGAACTAAAGAAATTTGGAGTCTCTGGGGCC 2728
Dh 5604 CTCTTTTACAGCCCTTATATACGCTGAGAACTAAAGAAATTTGGAGTCTCTGGGGCC 5663
Qy 2729 ATATATAAGAGCAGAAAGTGAAGTATATCATGTGTACTTTCAGAAATAGAGCCCTGCG 2788
Dh 5664 ATATATAAGAGCAGAAAGTGAAGTATATCATGTGTACTTTCAGAAATAGAGCCCTGCG 5723
Qy 2789 TCCCTATTCCTTATTCATGACCTTATTTCTTATGAGGAAGATCAGAGCAGAGAGAGA 2848
Dh 5724 TCCCTATTCCTTATTCATGACCTTATTTCTTATGAGGAAGATCAGAGCAGAGAGAGA 5783
Qy 2849 ACCTAGAAAAACCTTTGTCAAGCTTAATGAAACCAAACTTACTTTTGGAAATGCAACA 2908
Dh 5784 ACCTAGAAAAACCTTTGTCAAGCTTAATGAAACCAAACTTACTTTTGGAAATGCAACA 5843
Qy 2909 TCAATATGCAACCCACTTAAAGATGATTTGACTCAAGAGCTGGGCTTATTTCTGATGT 2968
Dh 5844 TCAATATGCAACCCACTTAAAGATGATTTGACTCAAGAGCTGGGCTTATTTCTGATGT 5903
Qy 2969 TGACCTGAAAAAAGATGTGACTCAGGCTGATTTGAACCCCTTCTGCTGCAACTAA 3028
Dh 5904 TGACCTGAAAAAAGATGTGACTCAGGCTGATTTGAACCCCTTCTGCTGCAACTAA 5963
Qy 3029 CACACTGAAACCTGCTCATGAGGAGACAGTGAAGTGAAGTATTTGCTGTTTTCAC 3088
Dh 5964 CACACTGAAACCTGCTCATGAGGAGACAGTGAAGTGAAGTATTTGCTGTTTTCAC 6023
Qy 3089 CATCTTTGATGAGACCAAAAGCTGTACTTCACTGAAATATGGAAGAACTGCAGGGC 3148
Dh 6024 CATCTTTGATGAGACCAAAAGCTGTACTTCACTGAAATATGGAAGAACTGCAGGGC 6083

QY	192	GGTGAAGCTGCTGTGGAGCGCAAGATTCTCTCTAGAGTGCCTAAATCTTTCCATTCAC	251
Db	121	GGTGAAGCTGCTGTGGAGCGCAAGATTCTCTCTAGAGTGCCTAAATCTTTCCATTCAC	180
QY	252	ACCTCAGTGTGTACAAAAAGACTGTGTTAGAAATTCAGGGTACGCTTTCACATAC	311
Db	181	ACCTCAGTGTGTACAAAAAGACTGTGTTAGAAATTCAGGGTACGCTTTCACATAC	240
QY	312	GCTAAGCCAAAGCCACCCTGATGGGTCTGCTAGAGTCTACCATCCAGCGTGAAGTTAT	371
Db	241	GCTAAGCCAAAGCCACCCTGATGGGTCTGCTAGAGTCTACCATCCAGCGTGAAGTTAT	300
QY	372	GATTAAGTGTGTATTACACTTAAGAATCGCTTCCATCTGTCACTTCACTGT	431
Db	301	GATTAAGTGTGTATTACACTTAAGAATCGCTTCCATCTGTCACTTCACTGT	360
QY	432	GGTGTATCTACTGTGAAGACTTCTGAGGAGCTGATATGTATGTACAGACATCAAG	491
Db	361	GGTGTATCTACTGTGAAGACTTCTGAGGAGCTGATATGTATGTACAGACATCAAG	420
QY	492	GAGAAAGAAAGATGATAAAGTCTTCCCTGTGTGAAGCCATACATATGCTGGAGCTGT	551
Db	421	GAGAAAGAAAGATGATAAAGTCTTCCCTGTGTGAAGCCATACATATGCTGGAGCTGT	480
QY	552	AAAGAGATGTGTCCATGCGCTCTGACCCACTGTGCTTACCTACTCATATCTTCTCAT	611
Db	481	AAAGAGATGTGTCCATGCGCTCTGACCCACTGTGCTTACCTACTCATATCTTCTCAT	540
QY	612	GTGAGACTGTGTAAAGACTTGAATTCAGGCGCTCATTTGAGCCCTACTAGATGTAAAGA	671
Db	541	GTGAGACTGTGTAAAGACTTGAATTCAGGCGCTCATTTGAGCCCTACTAGATGTAAAGA	600
QY	672	GGGAGTCTGGCCAAAGAAAAGACACAGACCTTGGACAAATTTATCTACTTTTGGTGA	731
Db	601	GGGAGTCTGGCCAAAGAAAAGACACAGACCTTGGACAAATTTATCTACTTTTGGTGA	660
QY	732	TTTGTATGAAGGAAAAGTTGGGCACTCAGAAAACAAAGACTCTTGATGCAAGATAGGAT	791
Db	661	TTTGTATGAAGGAAAAGTTGGGCACTCAGAAAACAAAGACTCTTGATGCAAGATAGGAT	720
QY	792	GCTGATCTGCTCGGGGCTGGGCTCTAAATGAGACACATGAATGATATGAACAGCTGT	851
Db	721	GCTGATCTGCTCGGGGCTGGGCTCTAAATGAGACACATGAATGATATGAACAGCTGT	780
QY	852	CTGCCAGGTCTGATTTGATGTCACAGAGAAATCAGTCTATTGGCATGTGATTTGAAATGGGC	911
Db	781	CTGCCAGGTCTGATTTGATGTCACAGAGAAATCAGTCTATTGGCATGTGATTTGAAATGGGC	840
QY	912	ACCACTGCTGAAGTGCATCAATTTCTGGAAGTGCACATTTCTTGTGAGAACCAT	971
Db	841	ACCACTGCTGAAGTGCATCAATTTCTGGAAGTGCACATTTCTTGTGAGAACCAT	900
QY	972	CGCCAGCGGTCTTGGAAATGTCCGCCATTAATCTTCTGCTGCTGAAACACGCTTGATG	1031
Db	901	CGCCAGCGGTCTTGGAAATGTCCGCCATTAATCTTCTGCTGCTGAAACACGCTTGATG	960
QY	1032	GACCTTGGACAGTTTCTACTGTTTGTATATCTCTTCCACCAACATATGAGATGGA	1091
Db	961	GACCTTGGACAGTTTCTACTGTTTGTATATCTCTTCCACCAACATATGAGATGGA	1020
QY	1092	GCTTATGTCAAAGTATGACAGCTGTGCCAGAGAAACCCCACTACGATGAAATATATGA	1151
Db	1021	GCTTATGTCAAAGTATGACAGCTGTGCCAGAGAAACCCCACTACGATGAAATATATGA	1080
QY	1152	GAGCGGAAAGACTATGATGATGATCTTACAGATTTGGAATGATGATGATGATGATGAT	1211
Db	1081	GAGCGGAAAGACTATGATGATGATCTTACAGATTTGGAATGATGATGATGATGATGAT	1140
QY	1212	GATGACAACTCTCTCTCTTATTCAAATTTGGCTAGTTGCCAAGAAGATCTTAAACT	1271
Db	1141	GATGACAACTCTCTCTCTTATTCAAATTTGGCTAGTTGCCAAGAAGATCTTAAACT	1200
QY	1272	TGGGTACATTACATTTGCTGCTGAAGAGAGAGACTGAGCATATGCTTCTTACCTCGCC	1331

Db	1201	TGTCACATTACATTCTGCTGCTGAAGAGAGAGACTGGGACTATGCTCCCTTAGCTCCGCGC	1261
Qy	1332	CCCGATGACAGAAAGTTATATAAAGTCAATATATTGAGCAATGGCCCTCAGCGGATTTGGTAGG	1391
Db	1261	CCCGATGACAGAAAGTTATATAAAGTCAATATTGAGCAATGGCCCTCAGCGGATTTGGTAGG	1321
Qy	1392	AAGTCAAAAAAGTCCGATTTATGGCATACACAGATGAACCTTTTAAGCTGTGAGCT	1451
Db	1321	AAGTCAAAAAAGTCCGATTTATGGCATACACAGATGAACCTTTTAAGCTGTGAGCT	1381
Qy	1452	ATTGACGATGATATAGGATCTTGGGACCTTATCTTATATGGGGAAGTTGGAGACACATG	1511
Db	1381	ATTGACGATGATATAGGATCTTGGGACCTTATCTTATATGGGGAAGTTGGAGACACATG	1441
Qy	1512	TTGATTTATTTTAAGATCAAGCAGACCATATPACATCTACCTCAGCGAAATCACT	1571
Db	1441	TTGATTTATTTTAAGATCAAGCAGACCATATPACATCTACCTCAGCGAAATCACT	1501
Qy	1572	GATGTCGCTCTTGTATATCAAGAGATTACCAAAAGTGTAACATTTGAAAGATTTT	1631
Db	1501	GATGTCGCTCTTGTATATCAAGAGATTACCAAAAGTGTAACATTTGAAAGATTTT	1561
Qy	1632	CCAAATTCGACAGAGAAATATCAATATAATTAATGAGCAGTGCATGTATGAAGATGGCCA	1691
Db	1561	CCAAATTCGACAGAGAAATATCAATATAATTAATGAGCAGTGCATGTATGAAGATGGCCA	1621
Qy	1692	ACTAATACAGATCCTCGTGCCGACCCGCTATTTACTTAGTTTGGTTAATATGAGAGGA	1751
Db	1621	ACTAATATAGATCCTCGTGCCGCTCAGCCCCCTATTTACTTAGTTTGGTTAATATGAGAGGA	1681
Qy	1752	GATCTAGCTTACAGACATATGGCCCTCTCTCATCTGTATACAAAAGAAATCTTAATACAA	1811
Db	1681	GATCTAGCTTACAGACATATGGCCCTCTCTCATCTGTATACAAAAGAAATTTGTATACAA	1741
Qy	1812	AGAGAAAACCAAGATATATGTACAGACAAAGAGAAATGTCATCTCTGTTTCTGTATTTGATGAG	1871
Db	1741	AGAGAAAACCAAGATATAATGTACAGACAAAGAGAAATGTCATCTCTGTTTCTGTATTTGATGAG	1801
Qy	1872	AACCGAATCTGTATCTCTCACAGAGAAATATCAAGCTTCTCCCAATCCAGCTGGAGTG	1931
Db	1801	AACCGAATCTGTATCTCTCACAGAGAAATATCAAGCTTCTCTCCCAATCCAGCTGGAGTG	1861
Qy	1932	CAGCTTGAGATCCAGAGTCCAAAGCCTCCAAATCATGTCACAGCATCAATGGCTATGT	1991
Db	1861	CAGCTTGAGATCCAGAGTCCAAAGCCTCCAAATCATGTCACAGCATCAATGGCTATGT	1921
Qy	1992	TTTGATAGTTTGAGTGTGTAGTTGTTGGACAGAGTGCGCATCTACTGTGATCTTAAAC	2051
Db	1921	TTTGATAGTTTGAGTGTGTAGTTGTTGGACAGAGTGTGATCTTAAAC	1981
Qy	2052	ATTGGAGACAGACTGACTTCTTCTGTTCTTCTCTGATATACCTTAAACACAAA	2111
Db	1981	ATTGGAGACAGACTGACTTCTTCTGTTCTTCTCTGATATACCTTAAACACAAA	2041
Qy	2112	ATTGTCATATGAAGACACACACACCCTATTTCCATTTCTCAGAGAAACGTCTTCATGTG	2171
Db	2041	ATTGTCATATGAAGACACACACACCCTATTTCCATTTCTCAGAGAAACGTCTTCATGTG	2101
Qy	2172	ATTGAAAACCCAGGTCTATGATTTTGGGGGTGCACAACTCAGATTTTGGGAAACGAGGC	2231
Db	2101	ATTGAAAACCCAGGTCTATGATTTTGGGGGTGCACAACTCAGATTTTGGGAAACGAGGC	2161
Qy	2232	ATTGACCGCTTACTGAAGTTTCTAGTTGTGACAAAGACACTGTGATATTATAGAGAC	2291
Db	2161	ATTGACCGCTTACTGAAGTTTCTAGTTGTGACAAAGACACTGTGATATTATAGAGAC	2221
Qy	2292	AGTTATAGATATTTACAGATATCTTGCGAGTAAAAAACAATGCAATTGAACAAAGAGC	2351
Db	2221	AGTTATAGATATTTTCAGCATCTTGCGAGTAAAAAACAATGCAATTGAACAAAGAGC	2281
Qy	2352	TTCTCCACGAA-----CCCAACA 2369	

Dh 2281 TTCTCCAGAAATTCAAGAGACCCCTAGCACTAGCAAAAGCAATTTATGCCACCCACCA 2340
Oy 2370 GTCTTGGAAACGCATCAACGGGAAATTACTGCTACTACTCTTCACTGATGAAGAA 2429
Dh 2341 GCTTGGAAACGCATCAACGGGAAATTACTGCTACTACTCTTCACTGATGAAGAA 2400
Oy 2430 ATTGACTATGATGATACCATATCACTGTAATGAAGAAAGATTTTGACATTTATGAT 2489
Dh 2401 ATTGACTATGATGATACCATATCACTGTAATGAAGAAAGATTTTGACATTTATGAT 2460
Oy 2490 GAGGATGAAAAATCAGAGCCCCCGCAGCTTTCAAAAGAAAAACGACACTATTTATGCT 2549
Dh 2461 GAGGATGAAAAATCAGAGCCCCCGCAGCTTTCAAAAGAAAAACGACACTATTTATGCT 2520
Oy 2550 GCAGTGGAGAGGCTGTGGATTTATGGATGAGTACTCCCAATGTTCTAAGAAACAG 2609
Dh 2521 GCAGTGGAGAGGCTGTGGATTTATGGATGAGTACTCCCAATGTTCTAAGAAAGAG 2580
Oy 2610 GCTCAGATGGGAGTGTCCCTAGTTCAAGAAAGTTGTTTCCAGGAATTTACTGATGCG 2669
Dh 2581 GCTCAGATGGGAGTGTCCCTAGTTCAAGAAAGTTGTTTCCAGGAATTTACTGATGCG 2640
Oy 2670 TCCTTACTCAGCCCTTATACCGTGGAGAACTAATGAACATTTGGGACTCCTGGGCCA 2729
Dh 2641 TCCTTACTCAGCCCTTATACCGTGGAGAACTAATGAACATTTGGGACTCCTGGGCCA 2700
Oy 2730 TATATAGAGCAAGATTGAGATTAATCATGTACTTTCAGAAATCAGGCTCTCGT 2789
Dh 2701 TATATAGAGCAAGATTGAGATTAATCATGTACTTTCAGAAATCAGGCTCTCGT 2760
Oy 2790 CCCATTCCCTTATTTAGCCCTTATTTCTTTAGGAAAGATCAGAGCAAGAGACAGAA 2849
Dh 2761 CCCATTCCCTTATTTAGCCCTTATTTCTTATGAGAAAGATCAGAGCAAGAGACAGAA 2820
Oy 2850 CCTAGAAAAAATTTGTCAAGCCCTAATGAAACCAAACTTACTTTGGAAATGCAACAT 2909
Dh 2821 CCTAGAAAAAATTTGTCAAGCCCTAATGAAACCAAACTTACTTTGGAAATGCAACAT 2880
Oy 2910 CATATGGCAACCCACTAAGATGATGTTGACTGCAAAAGCCTGGGCTTATTTCTGATGTT 2969
Dh 2881 CATATGGCAACCCACTAAGATGATGTTGACTGCAAAAGCCTGGGCTTATTTCTGATGTT 2940
Oy 2970 GACCGTGAAGAAAGATGTGCTACTCAGGCGCTGATGGAACCCCTGCTGCGCACACTAC 3029
Dh 2941 GACCGTGAAGAAAGATGTGCTACTCAGGCGCTGATGGAACCCCTGCTGCGCACACTAC 3000
Oy 3030 ACACCTGAACCCCTGCTCATGAGGAGACAAGTACAGAGATTTGCTGTTTTACCC 3089
Dh 3001 ACACCTGAACCCCTGCTCATGAGGAGACAAGTACAGAGATTTGCTGTTTTACCC 3060
Oy 3090 ATCTTTGATGAGACCAAAAGCTGTACTTCACTGAAATATGGAAGAAATGCAAGGCT 3149
Dh 3061 ATCTTTGATGAGACCAAAAGCTGTACTTCACTGAAATATGGAAGAAATGCAAGGCT 3120
Oy 3150 CCCCTCAATATCAATGGAAGATCCCACTTTTAAAGAAATATGCGTCCATGCAATC 3209
Dh 3121 CCCCTCAATATCAATGGAAGATCCCACTTTTAAAGAAATATGCGTCCATGCAATC 3180
Oy 3210 AATGCTACATATGATGATACACTACTGCTGCTAGTAATGGCTCAGAGTCAAGATGGA 3269
Dh 3181 AATGCTACATATGATGATGATACACTACTGCTGCTAGTAATGGCTCAGAGTCAAGATGGA 3240
Oy 3270 TGGTATGCTCTCAGCAGGCGCAATGAAACATCCATTTCTATTCATTTCACTGAGCAT 3329
Dh 3241 TGGTATGCTCTCAGCAGGCGCAATGAAACATCCATTTCTATTCATTTCACTGAGCAT 3300
Oy 3330 GGTTCACACTGTCGAAAAAAGAGAGATTAATAATGGCTTAATAATCTCATCAGGT 3389
Dh 3301 GGTTCACACTGTCGAAAAAAGAGAGATTAATAATGGCTTAATAATCTCATCAGGT 3360
Oy 3390 GTTTTGAACAGTGAAGATTTTACATCCTCAAGCTGGAATTTGGCGGGTGAATGCTT 3449
Dh 3361 GTTTTGAACAGTGAAGATTTTACATCCTCAAGCTGGAATTTGGCGGGTGAATGCTT 3420

Oy 3450 ATGGCGAGCATCTTACATGCTGGGATGAGCACACTTTTCTGGTGTACAGCAATTAAGTGT 3509
Dh 3421 ATGGCGAGCATCTTACATGCTGGGATGAGCACACTTTTCTGGTGTACAGCAATTAAGTGT 3480
Oy 3510 CAGACTCCCTGGGAATGGCTTCTGACACATTAAGATTTTTCAGATTAACGCTTCAGGA 3569
Dh 3481 CAGACTCCCTGGGAATGGCTTCTGACACATTAAGATTTTTCAGATTAACGCTTCAGGA 3540
Oy 3570 CAATATGACAGTGGGCCCAAGAGCTGGCCAGACTTCAATTAATCCGGATCAATCAATGCC 3629
Dh 3541 CAATATGACAGTGGGCCCAAGAGCTGGCCAGACTTCAATTAATCCGGATCAATCAATGCC 3600
Oy 3630 TGGACACCAAGAGACCCCTTTTCTTGATCAAGTGGATCTGTGGCACCAATGATTAAT 3689
Dh 3601 TGGACACCAAGAGACCCCTTTTCTTGATCAAGTGGATCTGTGGCACCAATGATTAAT 3660
Oy 3690 CACGCGATCAAGACCCAGAGGCGCGCTCAGAAAGTTCTCAGCCTCTACATCTCAGTTT 3749
Dh 3661 CACGCGATCAAGACCCAGAGGCGCGCTCAGAAAGTTCTCAGCCTCTACATCTCAGTTT 3720
Oy 3750 ATCATCATGATATGCTTTGATGAGGAAAGTGGCAGACTTATCGAGAAATTCACCTGGA 3809
Dh 3721 ATCATCATGATATGCTTTGATGAGGAAAGTGGCAGACTTATCGAGAAATTCACCTGGA 3780
Oy 3810 ACCTTAATGCTCTTTTGGCAATGTCATCTGGAATTAACACATATTTTAAAC 3869
Dh 3781 ACCTTAATGCTCTTTTGGCAATGTCATCTGGAATTAACACATATTTTAAAC 3840
Oy 3870 COTCAATTAATGCTCGATACATCCGTTTGCACCCCACTCAATTAAGATTCGCGACT 3929
Dh 3841 COTCAATTAATGCTCGATACATCCGTTTGCACCCCACTCAATTAAGATTCGCGACT 3900
Oy 3930 CTTCGATGAGATGATGGGCTGTGATTTAAATAGTTCCAGATGCCATTTGGGAATGAG 3989
Dh 3901 CTTCGATGAGATGATGGGCTGTGATTTAAATAGTTCCAGATGCCATTTGGGAATGAG 3960
Oy 3990 AGTAAAGCAATATCAGATGACAGATTAATGCTGCTCATCTTACCAATATGTTTGC 4049
Dh 3961 AGTAAAGCAATATCAGATGACAGATTAATGCTGCTCATCTTACCAATATGTTTGC 4020
Oy 4050 ACCGTGCTCCTTCAAAAGCTGACTCACTCCCAAGGAGAGTAATGCTCGGAGACT 4109
Dh 4021 ACCGTGCTCCTTCAAAAGCTGACTCACTCCCAAGGAGAGTAATGCTCGGAGACT 4080
Oy 4110 CAGTGAATTAATCCAAAGAGTGGCTGCAAGTGAAGTCTCCAGAAACATGAATGACACA 4169
Dh 4081 CAGTGAATTAATCCAAAGAGTGGCTGCAAGTGAAGTCTCCAGAAACATGAATGACACA 4140
Oy 4170 GGAGTAATCTACTCAGGAGTAATAATCTGCTTACACAGATGATGTAAGAGTTCCTC 4229
Dh 4141 GGAGTAATCTACTCAGGAGTAATAATCTGCTTACACAGATGATGTAAGAGTTCCTC 4200
Oy 4230 ATCTTCAGCAGTCAAGATGCGCATGATGAGTCTCTTTTTCAGAAATGGCAAGTAAG 4289
Dh 4201 ATCTTCAGCAGTCAAGATGCGCATGATGAGTCTCTTTTTCAGAAATGGCAAGTAAG 4260
Oy 4290 GTTTTCAAGGAAATCAAGACTCTTCAACACTGTTGGAATCTCTGACCCACCGTTA 4349
Dh 4261 GTTTTCAAGGAAATCAAGACTCTTCAACACTGTTGGAATCTCTGACCCACCGTTA 4320
Oy 4350 CTGACTCGCTACTTCAATTAACCCCAAGATGGGTGACACAGATGGCCCTGAGATG 4409
Dh 4321 CTGACTCGCTACTTCAATTAACCCCAAGATGGGTGACACAGATGGCCCTGAGATG 4380
Oy 4410 GAGGTTCTGGGCTGGAGGACAGACCTCTTACTAAGGCTGGCCACTGACAGCTTGCA 4469
Dh 4381 GAGGTTCTGGGCTGGAGGACAGACCTCTTACTAAGGCTGGCCACTGACAGCTTGCA 4440
Oy 4470 CTGCGTCACTCTCCCTCTCAGCTCCAGGCGCAATGCTCCCTGCTGGCTTCTAC 4529
Dh 4441 CTGCGTCACTCTCCCTCTCAGCTCCAGGCGCAATGCTCCCTGCTGGCTTCTAC 4500

QY	4530	CTTTGGCGTAAATCCACAGACACACGTCGCTTGAAGCTCTCTGAATATACATACATAGTC	4589
Db	4501	CTTTGGCGTAAATCCACAGACACACGTCGCTTGAAGCTCTCTGAATATACATACATAGTC	4560
UY	4590	CTGCATTTCTTGGTGGGGGGCCAGAGAGGTGCATCCAAATTAACTTAACCTTACCTAT	4649
Db	4561	CTGCATTTCTTGGTGGGGGGCCAGAGAGGTGCATCCAAATTAACTTAACCTTACCTAT	4620
QY	4650	TTTCTGCGAG 4658	
Db	4621	TTTCTGCGAG 4629	
RESULT	5		
ID	V23339		
AC	V23339	standard; DNA: 4670 BP.	
OS	17-AUG-1998	(first entry)	
OS	Human Factor-VIII gene lacking central B domain.		
OS	Factor-VIII; blood clotting; human; synthetic gene; codon usage; ss.		
OS	Homo sapiens.		
OS	Synthetic.		
PN	WO9812207-A1.		
PD	26-MAR-1998.		
PF	18-SEP-1997; U16639.		
PR	20-SEP-1996; US-717294.		
PA	(GHEO) GEN HOSPITAL CORP.		
PI	Haas J. Seed B:		
PI	WPI: 98-217200/19.		
PR	New synthetic eukaryotic gene(s) - in which non-preferred or less		
PT	preferred codon(s) are replaced to provide high level expression in		
PT	mammalian cell(s)		
PS	Example 3: Fig 12; 92pp; English.		
CC	This gene codes for a human Factor-VIII protein that lacks the		
CC	central B domain (amino acids 760-1639) of the native protein. In		
CC	a novel, claimed synthetic gene (see V23288), non-preferred or		
CC	less preferred codons of the native gene are replaced by codons		
CC	expressed by highly expressed human genes to provide high-level		
CC	expression in mammalian cells. The synthetic gene was assembled		
CC	from 29 pairs of oligonucleotides (see V23340-97) which served as		
CC	PCR templates. Synthetic genes of the invention (see also		
CC	V23289-91) are used for production of recombinant proteins in		
CC	mammalian cells at levels of at least 500% of those obtained using		
CC	the natural genes. They can also be used in gene therapy.		
SO	Sequence 4670 BP; 1325 A; 1072 C; 1036 G; 1237 T;		

	very Match	92.3%;	Score 4462;	DB 1;	Length 4670;
	Best Local Similarity	98.9%;	Pred. No. 0;		
	Matches 4520;	Conservative	0;	Mismatches 10;	Indels 42;
				Gaps	
QY	129	GCACACAGAAAGTACTACTGCTGGGTGCAGTGGGAACATCTCATATGGAGCATATATGCAAAGTAT	188		
Db	84	GCACACAGAAAGTACTACTGCTGGGTGCAGTGGGAACATCTCATATGGAGCATATATGCAAAGTAT	143		
QY	189	CTCGGTGAGCTCTCTGTGGACGCAGAATTTCCTCCCTAGATGGCCAAAACCTTTTCATTC	248		
Db	144	CTCGGTGAGCTCTCTGTGGACGCAGAATTTCCTCCCTAGATGGCCAAAACCTTTTCATTC	203		
QY	249	AACACCTCAGTCTGTACAAAAGACTCTGTTGTAGAAATTCAGGATACCTTTTCAAC	308		
Db	204	AACACCTCAGTCTGTACAAAAGACTCTGTTGTAGAAATTCAGGATACCTTTTCAAC	263		
QY	309	ATTCGCTAAGCCAAAGCCACCTCGATGGGTCTGCTAGAGTCTTACCATCCAGGCTGAGTT	368		
Db	264	ATTCGCTAAGCCAAAGCCACCTCGATGGGTCTGCTAGAGTCTTACCATCCAGGCTGAGTT	323		
QY	369	TATGATACAGTGTGTTATTTACCTTAAAGAACATAGGCTTCCCATCTGTCAGTCTTATGCT	428		
Db	324	TATGATACAGTGTGTTATTTACCTTAAAGAACATAGGCTTCCCATCTGTCAGTCTTATGCT	383		
QY	429	GTTGGTGTATCTACTATGGAAGCTTCTGAGGAGCTGAATATGATGATCAGACCAAGTCAA	488		

D	384	GTGGTGTATCTCTACTGGAAGAGCTTCGAGGAGAGCTGTAATATGATATATACAGCACTCAA	443
Q	489	AGGGAAGAAAGATGATATAAGTCTTCCTCGTGGAAGCCATACATATGCTCGGACGTC	548
D	444	AGGGAGAAAGAAAGATGATATAAGTCTTCCTCGTGGAAGCCATACATATGCTCGGACGTC	503
Q	549	CTGAAGAGAAATGGATGCAATGGCCTCGACCACTGGCCTTACCTACGATCATATCTTCT	608
D	504	CTGAAGAGAAATGGATGCAATGGCCTCGACCACTGGCCTTACCTACGATCATATCTTCT	563
Q	609	CATGTGACCTGGTAAAAAGACTTGAATTCAGGCCTCATTTGAGACCCCTACTAGTATGTAGA	668
D	564	CATGTGACCTGGTAAAAAGACTTGAATTCAGGCCTCATTTGAGACCCCTACTAGTATGTAGA	623
Q	669	GAAGGAGACTCGGCCCAAGGAAAAGAACACACACCTTGACAAATTTATCTACTTTTGGT	728
D	624	GAAGGAGACTCGGCCCAAGGAAAAGAACACACACCTTGACAAATTTATCTACTTTTGGT	683
Q	729	GTATTGTATGAAAGGAAAAGTTGGACATCCAGAAACAAAGACTCTTGATGCGAGATAG	788
D	684	GTATTGTATGAAAGGAAAAGTTGGACATCCAGAAACAAAGACTCTTGATGCGAGATAG	743
Q	789	GATGCTGATCTGCTCGGGCCTGGCCTAAATGCACACAGTCAGTGTATGTAAACAG	848
D	744	GATGCTGATCTGCTCGGGCCTGGCCTAAATGCACACAGTCAGTGTATGTAAACAG	803
Q	849	TCTCTGCCAGGCTGATTTGGATGCGCAAGAGAAATCACTATTTGGCATGATGATGGAATG	908
D	804	TCTCTGCCAGGCTGATTTGGATGCGCAAGAGAAATCACTATTTGGCATGATGATGGAATG	863
Q	909	GGCACACTCTCTGAAGTGCATCAATATTCCTCGAAGGTACACACATTTCTTGAGAAC	968
D	864	GGCACACTCTCTGAAGTGCATCAATATTCCTCGAAGGTACACACATTTCTTGAGAAC	923
Q	969	CATGCCACAGGCGCTTGGAATCTGGCCAAATCACTTCCTACAGCTCAACACTCTG	1028
D	924	CATGCCACAGGCGCTTGGAATCTGGCCAAATCACTTCCTACAGCTCAACACTCTG	983
Q	1029	ATGACCTTGGACAGTTTCTACTGTTTGTGCATATCTCTCCACCACATATGGCATTG	1088
D	984	ATGACCTTGGACAGTTTCTACTGTTTGTGCATATCTCTCCACCACATATGGCATTG	1043
Q	1089	GAAGCTATATGTCACAATGACAGACTGTGCCAGAGAAACCCCACTACGAATGAAAAATAT	1148
D	1044	GAAGCTATATGTCACAATGACAGACTGTGCCAGAGAAACCCCACTACGAATGAAAAATAT	1103
Q	1149	GAAAGAGGGAAGACATATGATGATATCTTACTGATTCGAATGGATGGTGGTCAGGTTT	1208
D	1104	GAAAGAGGGAAGACATATGATGATATCTTACTGATTCGAATGGATGGTGGTCAGGTTT	1163
Q	1209	GATGATGACACTCTCCTTCCTTTATCCAAATTCGCTCAGTTGCCAAGAAAGCATCTTAAA	1268
D	1164	GATGATGACAACCTCTCCTTCCTTTATCCAAATTCGCTCAGTTGCCAAGAAAGCATCTTAAA	1223
Q	1269	ACTTGGGACATTTCAATTTGGCGTGAAGAGAGAGATGGAGATATATGCTCCCTTACTCCTC	1328
D	1224	ACTTGGGACATTTCAATTTGGCGTGAAGAGAGAGATGGAGATATATGCTCCCTTACTCCTC	1283
Q	1329	GCCCCCGATGACAGAACTTTAAAAAGTCAATATTTGAACATGGCCCTCAGCGGAGTTGGT	1388
D	1284	GCCCCCGATGACAGAACTTTAAAAAGTCAATATTTGAACATGGCCCTCAGCGGAGTTGGT	1343
Q	1389	AGGAAGATCAAAAAAGCCGATTTATGGCATACACAGATGAAGAACTTTTAAGACTGTGAA	1448
D	1344	AGGAAGATCAAAAAAGCCGATTTATGGCATACACAGATGAAGAACTTTTAAGACTGTGAA	1403
Q	1449	GCTATTAGCATGATACAGAAATCTTGGACCTTTACTTTATGGGAAAGTTGAGACACA	1508
D	1404	GCTATTAGCATGATACAGAAATCTTGGACCTTTACTTTATGGGAAAGTTGAGACACA	1463
Q	1509	CTGTGTATTTATTTAAGATTCAGCAAGCAGACCATATATACATCTTACCTTCAGGAATC	1568
D	1464	CTGTGTATTTATTTAAGATTCAGCAAGCAGACCATATATACATCTTACCTTCAGGAATC	1523

1569 ACTGATGCCGTCCTTTGTATTCAGAGATTTACCAAGGTGTAAACATTTGAAGAT 1628
1524 ACTGATTCGGTCTCTTGTTATTCAGAGATTTACCAAGGTGTAAACATTTGAAGAT 1683
1629 TTTCCATTTGCGCAGAGAAATATTCAAATATTAATGACAGTGTAGAGATGGG 1688
1584 TTTCCATTTGCGCAGAGAAATATTCAAATATTAATGACAGTGTAGAGATGGG 1683
1689 CCAACTAATCAGATCCTCGGTGCTGACCCGCTATTACTCTAGTTTCGTTAATGAG 1748
1644 CCAACTAATCAGATCCTCGGTGCTGACCCGCTATTACTCTAGTTTCGTTAATGAG 1703
1749 AGAGATCTACTGAGACATGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1808
1704 AGAGATCTACTGAGACATGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1763
1809 CAAAGAGAAACAGATTAATGTCTCAGACAGAGATGTCAATCCTGTTTCTGTATTGAT 1868
1764 CAAAGAGAAACAGATTAATGTCTCAGACAGAGATGTCAATCCTGTTTCTGTATTGAT 1823
1869 GAGAACCGAAGCTGTGCTCTCAGAGAAATATACAGCTTTTCTCCCAATCCAGCTGGA 1928
1824 GAGAACCGAAGCTGTGCTCTCAGAGAAATATACAGCTTTTCTCCCAATCCAGCTGGA 1883
1929 GTGAGCTTGAAGATCCAGATTCAGAGCTCCCAACATCATGACAGATCAATGGCTAT 1988
1884 GTGAGCTTGAAGATCCAGATTCAGAGCTCCCAACATCATGACAGATCAATGGCTAT 1943
1989 GTTTTGTATGTTTGCAGTTGTCTGTTTGTGATGAGGTGGCATCTGTTACTCTCTA 2048
1944 GTTTTGTATGTTTGCAGTTGTCTGTTTGTGATGAGGTGGCATCTGTTACTCTCTA 2003
2049 AGCATTTGGACACAGACTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2108
2004 AGCATTTGGACACAGACTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2063
2109 AAAATGCTATGAAAGACACATCACCTATTCCTCATTCAGAGAAACGCTCTTATG 2168
2064 AAAATGCTATGAAAGACACATCACCTATTCCTCATTCAGAGAAACGCTCTTATG 2123
2169 TCGATGAAACCCAGGTCTATGATCTGGGTGGCCCAACTGAGACTTTCGGAACGA 2228
2124 TCGATGAAACCCAGGTCTATGATCTGGGTGGCCCAACTGAGACTTTCGGAACGA 2183
2229 GGCATGACCGCTTACTGAAAGTTTCTAGTTGTGACAAAGAACTGTTGATTATACGAG 2288
2184 GGCATGACCGCTTACTGAAAGTTTCTAGTTGTGACAAAGAACTGTTGATTATACGAG 2243
2289 GACAGTTATGAAGATATTTACGATCTGCTGAGTAAACAAATGCCATTGAACAAAGA 2348
2244 GACAGTTATGAAGATATTTACGATCTGCTGAGTAAACAAATGCCATTGAACAAAGA 2303
2349 AGCTTCTCCAGAA-----CCCA 2366
2304 AGCTTCTCCAGAA-----CCCA 2366
2367 CCACTTTGAAACCGCATCAACGGGAATACTGCTACTCTCTCTCTCTCTCTCTCTCTCT 2426
2364 CCACTTTGAAACCGCATCAACGGGAATACTGCTACTCTCTCTCTCTCTCTCTCTCTCT 2423
2427 GAATTTACTATGATGATACCATATCAGTTGAAATGAAGAAAGATTTGACATTTAT 2486
2424 GAATTTACTATGATGATACCATATCAGTTGAAATGAAGAAAGATTTGACATTTAT 2483
2487 GATGAGATGAATTCAGAGCCCGCAGCTTCAAAAAGAAACAGACATATTTAT 2546
2484 GATGAGATGAATTCAGAGCCCGCAGCTTCAAAAAGAAACAGACATATTTAT 2543
2547 GCTGAGATGAGAGCTCTGGGATTTAGGATGAGTGTAGCTCCCAATGTTCTAAGAAC 2606
2544 GCTGAGATGAGAGCTCTGGGATTTAGGATGAGTGTAGCTCCCAATGTTCTAAGAAC 2603

2607 AGGGCTCAGAGTGCGAGTGTCCCTCAGTTCAAGAAAGTTGTTTCCAGAAATTTACTGAT 2666
2604 AGGGCTCAGAGTGCGAGTGTCCCTCAGTTCAAGAAAGTTGTTTCCAGAAATTTACTGAT 2663
2667 GGGCTCTTTTACTAGCCCTTATACCGGTGAGAACTAAATGACATTTGGACTCTGGGG 2726
2664 GGGCTCTTTTACTAGCCCTTATACCGGTGAGAACTAAATGACATTTGGACTCTGGGG 2723
2727 CCAATATTAAGACAGAGTTGAAGATATATCATGTTAACTTTAGAAATCAGGCTCT 2786
2724 CCAATATTAAGACAGAGTTGAAGATATATCATGTTAACTTTAGAAATCAGGCTCT 2783
2787 GGTCCCTATCTCTTATTTTACCTTATTTCTATGAGAAATCAGAGCAAGAGCA 2846
2784 GGTCCCTATCTCTTATTTTACCTTATTTCTATGAGAAATCAGAGCAAGAGCA 2843
2847 GAACTGAGAAACCTTTGCAAGCCCTAATGAACCAAACTTACTTTGAAAGTCA 2906
2844 GAACTGAGAAACCTTTGCAAGCCCTAATGAACCAAACTTACTTTGAAAGTCA 2903
2907 CATCATATGCGACCCACTAAAGATGATTTGACTGCAAGCCCTGAGCTTATCTCTGAT 2966
2904 CATCATATGCGACCCACTAAAGATGATTTGACTGCAAGCCCTGAGCTTATCTCTGAT 2963
2967 GTTGACCTGAGAAAGATGTGCACTGAGGCTGATTTGAGACCTCTCTGTTCTGCACT 3026
2964 GTTGACCTGAGAAAGATGTGCACTGAGGCTGATTTGAGACCTCTCTGTTCTGCACT 3023
3027 AACACATGAACCTGCTCATGAGAGCAAGTGCAGTATGAGAAATGAGTCTGTTTTC 3086
3024 AACACATGAACCTGCTCATGAGAGCAAGTGCAGTATGAGAAATGAGTCTGTTTTC 3083
3087 ACCATCTTGTATGAGACCAAAAGCTGTACTCTAGTAAATATGAGAAACCTGAGG 3146
3084 ACCATCTTGTATGAGACCAAAAGCTGTACTCTAGTAAATATGAGAAACCTGAGG 3143
3147 GCTCCCTGCAATATCCAGATGAGAAAGATCCCACTTTTAAAGAAATTTCCCTTCCATGCA 3206
3144 GCTCCCTGCAATATCCAGATGAGAAAGATCCCACTTTTAAAGAAATTTCCCTTCCATGCA 3203
3207 ATCAATGGCTACATTAATGATGATACATCTGCTGCTTATGATGCTCAGATCAAGGAT 3266
3204 ATCAATGGCTACATTAATGATGATACATCTGCTGCTTATGATGCTCAGATCAAGGAT 3263
3267 CGATGTATCTGCTCAGCATGGGAGCAATGAAACATCATCTTATTCATTCAGTGA 3326
3264 CGATGTATCTGCTCAGCATGGGAGCAATGAAACATCATCTTATTCATTCAGTGA 3323
3327 CATGTCTTCACTGTACGAAAAAGAGAGATTAATGACACTGTACAAATCTCTATCCA 3386
3324 CATGTCTTCACTGTACGAAAAAGAGAGATTAATGACACTGTACAAATCTCTATCCA 3383
3387 GGTGTTTGTGAGACAGTGAAGATGTTACATCCAAACCTGGAATTTGGCGGGGGAATGC 3446
3384 GGTGTTTGTGAGACAGTGAAGATGTTACATCCAAACCTGGAATTTGGCGGGGGAATGC 3443
3447 CTATTTGGCAGCATCTACATGCTGGGATGAGCACACTTTTCTGTGTACAGCAATAG 3506
3444 CTATTTGGCAGCATCTACATGCTGGGATGAGCACACTTTTCTGTGTACAGCAATAG 3503
3507 TGTACAGTCTCCCTGGGAATGGCTTCTGACACATTTGAGATTTTTCAGATTTACGCTTCA 3566
3504 TGTACAGTCTCCCTGGGAATGGCTTCTGACACATTTGAGATTTTTCAGATTTACGCTTCA 3563
3567 GGAATATGAGACAGTGGGCCCCAAGCTGGCCAGACTTATTTATTCGGATCAATAT 3626
3564 GGAATATGAGACAGTGGGCCCCAAGCTGGCCAGACTTATTTATTCGGATCAATAT 3623
3627 GCTGAGACCAACGAGCCCTTTTCTTGATCAAGGTGATGCTGTGACCAATGAT 3686
3624 GCTGAGACCAACGAGCCCTTTTCTTGATCAAGGTGATGCTGTGACCAATGAT 3683
3687 ATTACGGCATCAAGACCCAGGGTCCCGTACAGAAAGTTCTCAGACCTCTACATCTCTCAG 3746

Db	3684	ATTTCAGGCGATTAAGACCCAGAGGAGGCCGTGACAGAAAGTTCTTCAGGCTCTACATCTCTCAG	37434
OY	3747	TTTATCATCATGTATAGTCTTGATGGAGAGAAGTGGCAGACTTATCGAGAGAAATTCAC	3806
Db	3744	TTTATCATCATGTATAGTCTTGATGGAGAGAAGTGGCAGACTTATCGAGAGAAATTCAC	3803
OY	3807	GGACCTTAATGTTCTTTGGCAATGTGGATTCATCTGGATTAACCAATATATTTT	3866
Db	3804	GGACCTTAATGTTCTTTGGCAATGTGGATTCATCTGGATTAACCAATATATTTT	3866
OY	3867	AACCTCCAAATATTTGCTCGATACATCCGTTTGACCCCAACTCATATTATAGATTGCAAC	3926
Db	3864	AACCTCCAAATATTTGCTCGATACATCCGTTTGACCCCAACTCATATTATAGATTGCAAC	3923
OY	3927	ACTCTTCGATGGAGTTGATGGGCTGTGATTTAAATAGTTGGACGATGGCATTTGGGAATG	3986
Db	3924	ACTCTTCGATGGAGTTGATGGGCTGTGATTTAAATAGTTGGACGATGGCATTTGGGAATG	3983
OY	3987	GAGAGTAAAGCAATATCAGATGACACATTTACTGTTCACTCACTTATACCAATATGTT	4046
Db	3984	GAGAGTAAAGCAATATCAGATGACACATTTACTGTTCACTCACTTATACCAATATGTT	4043
OY	4047	GCCACCTGCTCTCTTCAAAAGCTCGACTTCACCTCCAAAGGAGAGATATGCTGGAGA	4106
Db	4044	GCCACCTGCTCTCTTCAAAAGCTCGACTTCACCTCCAAAGGAGAGATATGCTGGAGA	4103
OY	4107	CCTCAGTGAATTAATCCAAAAGATGGCTCGAATGGAGACTTCAGAAAGCATGAAATC	4166
Db	4104	CCTCAGTGAATTAATCCAAAAGATGGCTCGAATGGAGACTTCAGAAAGCATGAAATC	4163
OY	4167	ACAGGATTAATCTACTCAGGGAGTAAATCTCTGCTTACAGCATGTATGTGAAGAGTTC	4226
Db	4164	ACAGGATTAATCTACTCAGGGAGTAAATCTCTGCTTACAGCATGTATGTGAAGAGTTC	4223
OY	4227	CTCATCTCCACAGTCGAAGATGGCCATGATGGACCTCTTTTTCAGAAATGGCAAGTA	4286
Db	4224	CTCATCTCCACAGTCGAAGATGGCCATGATGGACCTCTTTTTCAGAAATGGCAAGTA	4283
OY	4287	AAGGTTTTTCAGGGAAATCAAGACTCCTTCACACCTGTGTGTAACCTCTAGACCCACG	4346
Db	4284	AAGGTTTTTCAGGGAAATCAAGACTCCTTCACACCTGTGTGTAACCTCTAGACCCACG	4343
OY	4347	TTACGATCGCTCACTTCGAAATTCACCCCGAAGTGGGTGCACAGATTGCTGAGG	4406
Db	4344	TTACGATCGCTCACTTCGAAATTCACCCCGAAGTGGGTGCACAGATTGCTGAGG	4403
OY	4407	ATGAGATTTGCGGCTCGAGAGCAGAGACCTTACTGAGGGTGGCCACTGAGACCTG	4466
Db	4404	ATGAGATTTGCGGCTCGAGAGCAGAGACCTTACTGAGGGTGGCCACTGAGACCTG	4463
OY	4467	CCACTGCCGTACTCTTCCTCTCAGACTTCAGAGGAGTGTCCCTCGCTGGCTTGCTTC	4526
Db	4464	CCACTGCCGTACTCTTCCTCTCAGACTTCAGAGGAGTGTCCCTCGCTGGCTTGCTTC	4523
OY	4527	TACCTTGTGTAAATCTCTGAGAGACACGCTTGAAAGCTTCCTCAATTAATCATCA	4586
Db	4524	TACCTTGTGTAAATCTCTGAGAGACACGCTTGAAAGCTTCCTCAATTAATCATCA	4583
OY	4587	GTCCTGATTTCTTTGGTGGGGGCCAGAGAGGTGCATCAATTAATCACTTAACCTTACC	4646
Db	4584	GTCCTGATTTCTTTGGTGGGGGCCAGAGAGGTGCATCAATTAATCACTTAACCTTACC	4643
OY	4647	TATTTTCTGCAG 4658	
Db	4644	GTGACCTGCAG 4655	
RESULT 6			
NB1544			
ID NB1544 standard: DNA: 4830 BP.			
NC NB1544;			
DT 04-DEC-1990 (first entry)			

DE Human Factor VIII:C analog having exon 14 deleted.
KW Human Factor VIII:C analog; exon deletion; coagulation disorders;
KW haemophilia; ss.
PN EP-265778-A.
PD 04-MAY-1988.
PF 14-OCT-1987; 115043.
PR 15-OCT-1986; US-919153.
PA (RORE) Rorer Int Overseas.
PI Sarver N, Drohan W.;
PI WPI; 88-120930/18.
PR Human Factor VIII:C analogue free of other proteins - produced from
PI recombinant DNA and used in the treatment of coagulation disorders.
PS Claim 3; Page 13; 42pp. English.
CC The genetically engineered analogue can provide a dependable and
CC readily available therapeutic agent to be used in the treatment of
CC haemophilia and coagulation disorders in humans or animals.
CC See also N81543-45.
SQ Sequence 4830 BP; 1392 A; 1091 C; 1058 G; 1289 T;

Query Match	Best Local Similarity	89.5%	Score 4322.4	DB 1	Length 4830
Query Match	Best Local Similarity	94.8%	Pred. No. 0	Mismatches 4579	Conservative 0
Matches 4579	Conservative	0	Mismatches 11	Indels 240	Gaps 1
QY	72	ATGCATAATAGACCTCTCCACACCTGCTTCTTCTGTCCTTTGTCGATTCGCTTTGATGCC	131		
Db	1	ATGCATAATAGACCTCTCCACACCTGCTTCTTCTGTCCTTTGTCGATTCGCTTTGATGCC	60		
QY	132	ACCAAGATATCTACTCTGGGTGACAGTGAAGTGTGATATATGCAAGTATCTC	191		
Db	61	ACCAAGATATCTACTCTGGGTGACAGTGAAGTGTGATATATGCAAGTATCTC	120		
QY	192	GGTGTGTCGCCGTGGAGGCAAGATTTCTCTTAAGTGGCAAAATCTTTTCCATTTCAAC	251		
Db	121	GGTGTGTCGCCGTGGAGGCAAGATTTCTCTTAAGTGGCAAAATCTTTTCCATTTCAAC	180		
QY	252	ACCTCAGTCGTGTACAAAAGACCTCTGTTTGTAGAAATTCACGGATCACCTTTTCAACATC	311		
Db	181	ACCTCAGTCGTGTACAAAAGACCTCTGTTTGTAGAAATTCACGGATCACCTTTTCAACATC	240		
QY	312	GCTAAGCCAAAGGCCACCTGGATGGGTGCTGTAAGTCTTACCATCCAGGCTGAGTTTAT	371		
Db	241	GCTAAGCCAAAGGCCACCTGGATGGGTGCTGTAAGTCTTACCATCCAGGCTGAGTTTAT	300		
QY	372	GATACAGTGTCTATTACTTAAAGAACTTTGGTTCCATCCGTGTCAGTCTCATCTCTT	431		
Db	301	GATACAGTGTCTATTACTTAAAGAACTTTGGTTCCATCCGTGTCAGTCTCATCTCTT	360		
QY	432	GGTGTATCTACTGGAAAGCTTCTGAGGAGCTGAATATGATGATCAACAGTCAAAAG	491		
Db	361	GGTGTATCTACTGGAAAGCTTCTGAGGAGCTGAATATGATGATCAACAGTCAAAAG	420		
QY	492	GAGAAAGAAATGATTAAGTCTTCCCTGTGTGAAAGCCATACATATGTCTGCAAGTCTTG	551		
Db	421	GAGAAAGAAATGATTAAGTCTTCCCTGTGTGAAAGCCATACATATGTCTGCAAGTCTTG	480		
QY	552	AAAAGAAATGATTCGAATGGCTGTGACCCACTGTGACCTTACTCTCTCATATCTTTCTCAT	611		
Db	481	AAAAGAAATGATTCGAATGGCTGTGACCCACTGTGACCTTACTCTCTCATATCTTTCTCAT	540		
QY	612	GTGACCTGTGTAAGAAAGCTTGAATTCAGGCTCTCATTTGGAGCCCTACTAGTATGTAGAA	671		
Db	541	GTGACCTGTGTAAGAAAGCTTGAATTCAGGCTCTCATTTGGAGCCCTACTAGTATGTAGAA	600		
QY	672	GGGAGTCTGGCCAAAGGAAAGACACAGACCTTGCACAAATTTATACTACTTTTGTCTGTA	731		
Db	601	GGGAGTCTGGCCAAAGGAAAGACACAGACCTTGCACAAATTTATACTACTTTTGTCTGTA	660		
QY	732	TTTATGAAAGGAAAGTGTGCACTCAAGAAACAAAGAACTCTGATAGCAGATAGGAT	791		
Db	661	TTTATGAAAGGAAAGTGTGCACTCAAGAAACAAAGAACTCTGATAGCAGATAGGAT	720		
QY	792	GCTGCATCTGCTGGGCTGCTTAAATGACACAGTCAATGTTATGTAAACAGGCT	851		

Db 721 GGTGATCTGCTGGGCTGGGCTTAATATGACACAGTCATAGTATATGTAACAGGTCT 780
Qy 852 CTGCCAGGTCTGTATGGATGCCACAGGAATCAGTCTATGGCAGTATGGATGGC 911
Db 781 CTGCCAGGTCTGTATGGATGCCACAGGAATCAGTCTATGGCAGTATGGATGGC 840
512 ACCACTCTGAAAGTGCACATATCTCCGAAAGTGCACATCTTCTTGAGGAACCAT 971
Db 841 ACCACTCTGAAAGTGCACATATCTCCGAAAGTGCACATCTTCTTGAGGAACCAT 900
Qy 972 CGCCAGGCTCTGTGGAATCTCGCAATACCTTCTACTGCTCAACACCTCTGAT 1031
Db 901 CGCCAGGCTCTGTGGAATCTCGCAATACCTTCTACTGCTCAACACCTCTGAT 960
Qy 1032 GACCTTGACAGTCTCTACTGTTTGTATCTCTCCACACATGATGGATGAA 1091
Db 961 GACCTTGACAGTCTCTACTGTTTGTATCTCTCCACACATGATGGATGAA 1020
Qy 1092 GCTTATGTCAAGTATGACAGTCTGTCAGAGAACCCCACTCGATGAAAAATATGAA 1151
1021 GCTTATGTCAAGTATGACAGTCTGTCAGAGAACCCCACTCGATGAAAAATATGAA 1080
Qy 1152 GAAGCGGAAGACTATGATGATGATCTTACTGATTTCTGAATGGATGGTCAAGTTGAT 1211
Db 1081 GAAGCGGAAGACTATGATGATGATCTTACTGATTTCTGAATGGATGGTCAAGTTGAT 1140
Qy 1212 GATGACACTCTCTCTCTTATCCAAATTCGCTCAGTTGCCAAGACATCCTAAACT 1271
Db 1141 GATGACACTCTCTCTCTTATCCAAATTCGCTCAGTTGCCAAGACATCCTAAACT 1200
Qy 1272 TGGGTACATTAATGCTGCTGTAAGAGAGAGTGGAGTATGCTGCTTACCTGCTGCT 1331
Db 1201 TGGGTACATTAATGCTGCTGTAAGAGAGAGTGGAGTATGCTGCTTACCTGCTGCT 1260
Qy 1332 CCCGATGACAGAGTATATATAAGTCAATATTTGAACAATGGCCCTCAGCGATTTGATAG 1391
Db 1261 CCCGATGACAGAGTATATATAAGTCAATATTTGAACAATGGCCCTCAGCGATTTGATAG 1320
Qy 1392 AAGTACAAAAGTCCGATTTATGCAATACAGATGAAACCTTTAAGACTCGTGAAGT 1451
Db 1321 AAGTACAAAAGTCCGATTTATGCAATACAGATGAAACCTTTAAGACTCGTGAAGT 1380
Qy 1452 ATTCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1511
Db 1381 ATTCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
Qy 1512 TTGATTTATTTAAGATCAAGCAAGCAGATATACATACCTCAGCGAATCAGT 1571
Db 1441 TTGATTTATTTAAGATCAAGCAAGCAGATATACATACCTCAGCGAATCAGT 1500
Qy 1572 GATGTCGCTCTTTGATTTCAAGAGATTAACAAAAGGTGTAAACATTTGAAGATTTT 1631
Db 1501 GATGTCGCTCTTTGATTTCAAGAGATTAACAAAAGGTGTAAACATTTGAAGATTTT 1560
Qy 1632 CCAATTTCTCCAGAGAAATATTTCAAAATTAATGAGAGTGTAGAAAGTGGCCA 1691
Db 1561 CCAATTTCTCCAGAGAAATATTTCAAAATTAATGAGAGTGTAGAAAGTGGCCA 1620
Qy 1692 ACTAAATCAGATCTGCTGGTGGCTGACCGCTATTAATGATTTGTTAATATGAGAG 1751
Db 1621 ACTAAATCAGATCTGCTGGTGGCTGACCGCTATTAATGATTTGTTAATATGAGAG 1680
Qy 1752 GATCTAGCTCAGAGACTCATTTGGCCCTCTCTCATCTGCTACAAAGAACTGTAGATCA 1811
Db 1681 GATCTAGCTCAGAGACTCATTTGGCCCTCTCTCATCTGCTACAAAGAACTGTAGATCA 1740
Qy 1812 AAGAGAAACCATATATGTCAGCAAGAGAAATGATCTCTGTTTCTGTATTTGATGAG 1871
Db 1741 AAGAGAAACCATATATGTCAGCAAGAGAAATGATCTCTGTTTCTGTATTTGATGAG 1800
Qy 1872 AACCGAAGCTGATCTCAGAGAAATATACAGGCTTTCTCCCATCAGCTGAGTGG 1931
Db 1801 AACCGAAGCTGATCTCAGAGAAATATACAGGCTTTCTCCCATCAGCTGAGTGG 1860
Qy 1932 CAGCTTGAGGATTCAGAGTTCCAAAGCTCCAAATCATCATGACAGCATCAATGGCTATGTT 1991
Db 1861 CAGCTTGAGGATTCAGAGTTCCAAAGCTCCAAATCATCATGACAGCATCAATGGCTATGTT 1920
Qy 1992 TTTGATAGTTTGCAGTTGTCAGTTGTTGTCAGTGGTGGCATGAGTCAATCTCAAGC 2051
Db 1921 TTTGATAGTTTGCAGTTGTCAGTTGTTGTCAGTGGTGGCATGAGTCAATCTCAAGC 1980
Qy 2052 ATTGAGACACAGACTGACT 2111
Db 1981 ATTGAGACACAGACTGACT 2040
Qy 2112 ATGCTTATGAAGACACACTACCTATTCATCTCAGAGAAACTCTTCAATGTCG 2171
Db 2041 ATGCTTATGAAGACACACTACCTATTCATCTCAGAGAAACTCTTCAATGTCG 2100
Qy 2172 ATGGAACACCCAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2231
Db 2101 ATGGAACACCCAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2160
Qy 2232 ATGACCGCTTACTGAGGTTTCTAGTTGACAAAGACACTGTTGATTTACAGAGAC 2291
Db 2161 ATGACCGCTTACTGAGGTTTCTAGTTGACAAAGACACTGTTGATTTACAGAGAC 2220
Qy 2292 AGTTATGAAGATTTTACACATCTCTGAGTAAAAACAATGCCATTTACACAGAGC 2351
Db 2221 AGTTATGAAGATTTTACACATCTCTGAGTAAAAACAATGCCATTTACACAGAGC 2280
Qy 2352 TTCTCCAG----- 2361
Db 2281 TTCTCCAGAAATTCATATTTGATCTCTGTTGGATTAACCATGATGATCTAGATA 2340
Qy 2361 ----- 2361
Db 2341 CCAAAAGAGAGTGAATCCCAAGAGAGTCAACGAAAAAACACTTTTAAGAAAAAG 2400
Qy 2361 ----- 2361
Db 2401 GATACATTTTGTCTGTAAGCTTTGTAAGCAATCATGATAGCAATTAATGAG 2460
Qy 2361 ----- 2361
Db 2461 GGAACAATTAAGCCGAAATAGAGTCACTGGGACACAGAGTGAAGTGAAGCTG 2520
Qy 2361 -----AACCACAGCTTTGAAGCCCATCAAGGGAATACTGTAATCTT 2411
Db 2521 TGCTCAAAACCCACAGCTCTTGAAGCCCATCAAGGGAATACTGTAATCTT 2580
Qy 2412 CAGTCAGATCAAGAGAAATGACATGATATACCATATCAGTTGAAATGAAGAGAA 2471
Db 2581 CAGTCAGATCAAGAGAAATGACATGATATACCATATCAGTTGAAATGAAGAGAA 2640
Qy 2472 GATTTGACATTTATATAGATGATAAATCAGAGCCCCCGCAGCTTTCAAAAAGAAACA 2531
Db 2641 GATTTGACATTTATATAGATGATAAATCAGAGCCCCCGCAGCTTTCAAAAAGAAACA 2700
Qy 2532 CGACACTATTTATGCTCAGTGAAGAGCTCTGGGATTTATGAGTGAAGTCTCCCA 2591
Db 2701 CGACACTATTTATGCTCAGTGAAGAGCTCTGGGATTTATGAGTGAAGTCTCCCA 2760
Qy 2592 CATGTTCTAAGAAACAGGCTCAGAGTGAAGTCTCAGTTCAAGAAAGTTGTTTC 2651
Db 2761 CATGTTCTAAGAAACAGGCTCAGAGTGAAGTCTCAGTTCAAGAAAGTTGTTTC 2820
Qy 2652 CAGGAATTTACGATGCTCTTACTCAGCCCTTATACCTGAGACATTAATGAACAT 2711
Db 2821 CAGGAATTTACGATGCTCTTACTCAGCCCTTATACCTGAGACATTAATGAACAT 2880
Qy 2712 TTGGAGCTCTGAGGCTCATATATTAAGAGAGAAAGTGAAGATTAATCATGTAATTTTC 2771
Db 2881 TTGGAGCTCTGAGGCTCATATATTAAGAGAGAAAGTGAAGATTAATCATGTAATTTTC 2940

QY 2772 AGAATCAGGCTCTGCTCCCTATTCCTTCTATCTTACCTTTATTTCTTATGAGAGAT 2831
| | | | |
Db 2941 AGAATTCAGGCTCTGCTCCCTATTCCTTCTATCTTACCTTTATTTCTTATGAGAGAT 3000
| | | | |
QY 2932 CAGAGCAGGAGCAGAACCTAGAAAAAACTTTGTCAGACCTTAATGAACCAAACTTAC 2891
| | | | |
Db 3001 CAGAGCAGGAGCAGAACCTAGAAAAAACTTTGTCAGACCTTAATGAACCAAACTTAC 3060
| | | | |
QY 2892 TTTTGGAAAGTCACATCATATGACCCCACTAAAGATGAGTTGACTGCAAAAGCCTG 2951
| | | | |
Db 3061 TTTTGGAAAGTCACATCATATGACCCCACTAAAGATGAGTTGACTGCAAAAGCCTG 3120
| | | | |
QY 2952 GCTTATTTCTGATGATGAGCTGGAAAAAGATGCACTAGGCTGATGAGCCCTT 3011
| | | | |
Db 3121 GCTTATTTCTGATGATGAGCTGGAAAAAGATGCACTAGGCTGATGAGCCCTT 3180
| | | | |
QY 3012 CTGCTCTGCACTAATACACTGAAACCTGCTCATGAGGACAAAGTACAGTACAGAA 3071
| | | | |
Db 3181 CTGCTCTGCACTAATACACTGAAACCTGCTCATGAGGACAAAGTACAGTACAGAA 3240
| | | | |
QY 3072 TTTGCTCTGTTTTCACCATCTTGTATGAGACCAAAAGCTGTACTCTGAAATATG 3131
| | | | |
Db 3241 TTTGCTCTGTTTTCACCATCTTGTATGAGACCAAAAGCTGTACTCTGAAATATG 3300
| | | | |
QY 3132 GAAAGAACTGCAGGCTCCCTGCAATATCCAGATGGAAATGCCACTTTTAAAGAAAT 3191
| | | | |
Db 3301 GAAAGAACTGCAGGCTCCCTGCAATATCCAGATGGAAATGCCACTTTTAAAGAAAT 3360
| | | | |
QY 3192 TATGCTTCATGCAATCAATGAGTACATATATGATACACTGCTTATGATGCT 3251
| | | | |
Db 3361 TATGCTTCATGCAATCAATGAGTACATATATGATACACTGCTTATGATGCT 3420
| | | | |
QY 3252 CAGATTAAGGATTCATGATGATCTGTCAGCATGGGACCAATGAAACATTCATCT 3311
| | | | |
Db 3421 CAGATTAAGGATTCATGATGATCTGTCAGCATGGGACCAATGAAACATTCATCT 3480
| | | | |
QY 3312 ATTCAATTCAGTGCATGCTGTCTGCTGAGAAAAAGAGAGTAAATAAGGACCTG 3371
| | | | |
Db 3481 ATTCAATTCAGTGCATGCTGTCTGCTGAGAAAAAGAGAGTAAATAAGGACCTG 3540
| | | | |
QY 3372 TACATCTCTATCCAGGTGTTTTGAGACAGTGGAAATGTTACCATCAAGCTGGAAT 3431
| | | | |
Db 3541 TACATCTCTATCCAGGTGTTTTGAGACAGTGGAAATGTTACCATCAAGCTGGAAT 3600
| | | | |
QY 3432 TGGGAGGTGGAAATGCTTATTTGGGAGCATCTATGCTGGATGAGACACATTTTCTG 3491
| | | | |
Db 3601 TGGGAGGTGGAAATGCTTATTTGGGAGCATCTATGCTGGATGAGACACATTTTCTG 3660
| | | | |
QY 3492 GTGTACGACATATAGTGTACAGTCCCTGGGAATGGCTTCTGGACATTAAGATTTT 3551
| | | | |
Db 3661 GTGTACGACATATAGTGTACAGTCCCTGGGAATGGCTTCTGGACATTAAGATTTT 3720
| | | | |
QY 3552 CAATTTACACTTCAGACCAATATGACAGTGGGCCCAAAAGCTGGCCAGACTTCATTA 3611
| | | | |
Db 3721 CAATTTACACTTCAGACCAATATGACAGTGGGCCCAAAAGCTGGCCAGACTTCATTA 3780
| | | | |
QY 3612 TCCGATCATCATGATGCTGAGACCAAGAGCCCTTTCTTGGATGCAAGGTGATCTG 3671
| | | | |
Db 3781 TCCGATCATCATGATGCTGAGACCAAGAGCCCTTTCTTGGATGCAAGGTGATCTG 3840
| | | | |
QY 3672 TTGGACCAATGATTTATTCAGGCAATCAAGCCAGGCTCCGTCGAGAGTTCTCCAG 3731
| | | | |
Db 3841 TTGGACCAATGATTTATTCAGGCAATCAAGCCAGGCTCCGTCGAGAGTTCTCCAG 3900
| | | | |
QY 3732 CTCTAATCTCTAGTTTATCATCATGATATAGTCTTGAGGAGAGAGTGGACACTTAT 3791
| | | | |
Db 3901 CTCTAATCTCTAGTTTATCATCATGATATAGTCTTGAGGAGAGAGTGGACACTTAT 3960
| | | | |
QY 3792 CGAGGAAATTCACCTGGAACCTTAATGCTTCTTGGCAATGATGATTCATCTGGGATA 3851
| | | | |
Db 3961 CGAGGAAATTCACCTGGAACCTTAATGCTTCTTGGCAATGATGATTCATCTGGGATA 4020
| | | | |

QY 3852 AACACAAATTTTAAACCTCCAAATTAATGTCGATACATCCGTTTGCACCCCACTCAT 3911
| | | | |
Db 4021 AACACAAATTTTAAACCTCCAAATTAATGTCGATACATCCGTTTGCACCCCACTCAT 4080
| | | | |
QY 3912 TATGAGCTTGGCAGCACTCTTCCGCAATGAGATGATGGGCTGTATTAATTAATTGCAGC 3971
| | | | |
Db 4081 TATGAGCTTGGCAGCACTCTTCCGCAATGAGATGATGGGCTGTATTAATTAATTGCAGC 4140
| | | | |
QY 3972 ATGCCATTGGGAATGAGAGATTAAGCAATATCATGATGACAGATTCGCTTATCTCAT 4031
| | | | |
Db 4141 ATGCCATTGGGAATGAGAGATTAAGCAATATCATGATGACAGATTCGCTTATCTCAT 4200
| | | | |
QY 4032 TTTTACCAATATGTTTGGCACCCTGCTCTTCAAAAGCTGCACCTTCAAGGAGAG 4091
| | | | |
Db 4201 TTTTACCAATATGTTTGGCACCCTGCTCTTCAAAAGCTGCACCTTCAAGGAGAG 4260
| | | | |
QY 4092 AGTAATGCTGGAGACCTCGAGTGAATTAATCCAAAGAGGCTGCAAGGACCTTCAG 4151
| | | | |
Db 4261 AGTAATGCTGGAGACCTCGAGTGAATTAATCCAAAGAGGCTGCAAGGACCTTCAG 4320
| | | | |
QY 4152 AAGACATGAAGTCAAGAGATTAATCACTACTCAGGAGATTAATCTGCTTACCAAGATG 4211
| | | | |
Db 4321 AAGACATGAAGTCAAGAGATTAATCACTACTCAGGAGATTAATCTGCTTACCAAGATG 4380
| | | | |
QY 4212 TATGTAAGAGATTCCTCATCTCAGCAGTCAAGATGSCCATCATGACCTCTTTT 4271
| | | | |
Db 4381 TATGTAAGAGATTCCTCATCTCAGCAGTCAAGATGSCCATCATGACCTCTTTT 4440
| | | | |
QY 4272 CAGATGGCAAAATGAAGTTTTCAGGGAATCAAGACTCCCTCAGCTGAGTGAAC 4331
| | | | |
Db 4441 CAGATGGCAAAATGAAGTTTTCAGGGAATCAAGACTCCCTCAGCTGAGTGAAC 4500
| | | | |
QY 4332 TCTCTAGACCCACCTTACTGACTGCTACTCTGCAATTCACCCCAAGATTGGGTGAC 4391
| | | | |
Db 4501 TCTCTAGACCCACCTTACTGACTGCTACTCTGCAATTCACCCCAAGATTGGGTGAC 4560
| | | | |
QY 4392 CAGATGGCCCTGAGATGAGGTTCTGGGCTGCGAGGCAAGAGACTTACTGAGGTTG 4451
| | | | |
Db 4561 CAGATGGCCCTGAGATGAGGTTCTGGGCTGCGAGGCAAGAGACTTACTGAGGTTG 4620
| | | | |
QY 4452 CCACCTGACACCTGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4511
| | | | |
Db 4621 CCACCTGACACCTGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4680
| | | | |
QY 4512 CCTGCTGCTGCTTCTACTTGTGCTAAATCTGACAGACACTGCTGGAAGCTCTG 4571
| | | | |
Db 4681 CCTGCTGCTGCTTCTACTTGTGCTAAATCTGACAGACACTGCTGGAAGCTCTG 4740
| | | | |
QY 4572 AATTACTATCATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4631
| | | | |
Db 4741 AATTACTATCATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4800
| | | | |
QY 4632 AACTTAATCTTACCTATTTCTGCAGCTG 4661
| | | | |
Db 4801 AACTTAATCTTACCTATTTCTGCAGCTG 4830
| | | | |
RESULT 7
ID N90654 standard; DNA; 4275 BP.
AC N90654;
DT 26-JUN-1990 (first entry)
DE DNA encoding 740 Arg-1649 Glu human Factor VIII:C
KW Human Factor VIII:C; Ad.RE.neo; 740 Arg-1649 Glu Factor VIII:C;
KW haemophilia A.
OS Homo sapiens.
FH Key
FT cds Location/Qualifiers
FT 1..4275
FT EP-306968-A.
PD 15-MAR-1989.
PF 09-SEP-1988; 114769.
PR 08-APR-1988; JP-085454.

PA (KAGA) Chemo-Sero-Therap (Telj).
 PI Sugiyama T, Masuda K, Tajima Y, Yonemura H:
 DR WPI, 89-078467/11.
 DR P-PSDB; P91165.
 PT Prod. of recombinant human Factor-VIII-C -
 PT using animal cells transformed with a vector contg. the gene for
 PS Factor VIII:C and a promoter
 PS Figure 1(1) - 1(13); : 32p; English.
 CC When translated, Arg-740 of the carboxyl terminus of the H chain is
 CC directly bonded by a peptide bond to Glu-1649 of the amino terminus of
 CC L chain. It is used to transform animal cells so that they produce
 CC human Factor VIII:C. A prefd. expression vector is Plasmid Ad.RE.neo.
 CC The expression vector has at least one promoter upstream of nt0654.
 CC The transformants can constantly and continuously produce human Factor
 CC VIII:C in high yield on a commercial scale. The human Factor VIII:C so
 CC produced is considered to corresp. to the smallest species of active and
 CC intact Factor VIII:C molecules in the human blood plasma. It is useful
 CC for treating haemophilia A patients.
 SQ Sequence 4275 BP; 1245 A; 941 C; 945 G; 1144 T;

Query Match 87.3%; Score 4216.6; DB 1; Length 4275;
 Best Local Similarity 98.9%; Pred. No. 0;
 Matches 4271; Conservative 0; Mismatches 4; Indels 42; Gaps 1;

QY 129 GCCACGAGAGTACTACTCTGGGCTGAGCTGAGTCTGATGGGACTATATGCAAGAT 188
 DB 1 GCCACGAGAGTACTACTCTGGGCTGAGCTGAGTCTGATGGGACTATATGCAAGAT 60
 QY 189 CTGGGTAGCTGCTCTGAGAGCAAGATTCTCTAGAGTGCAGAAATCTTTCCATTC 248
 DB 61 CTGGGTAGCTGCTCTGAGAGCAAGATTCTCTAGAGTGCAGAAATCTTTCCATTC 120
 QY 249 AACACCTAGCTGCTGACAAAAGACTCTGTTTGTAGAAATTCACGATCACCTTTCAAC 308
 DB 121 AACACCTAGCTGCTGACAAAAGACTCTGTTTGTAGAAATTCACGATCACCTTTCAAC 180
 QY 309 ATGCTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 368
 DB 181 ATGCTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
 QY 369 TATGATACAGTGTCTTACCTTAAGAAACATGAGTCCCTGAGTCCATGCTCATGCT 428
 DB 241 TATGATACAGTGTCTTACCTTAAGAAACATGAGTCCCTGAGTCCATGCTCATGCT 300
 QY 429 GTTGGTGTATCTACTGAGAAAGCTTCTGAGGAGCTGAAATGATGATCAGACCACTCAA 488
 DB 301 GTTGGTGTATCTACTGAGAAAGCTTCTGAGGAGCTGAAATGATGATCAGACCACTCAA 360
 QY 489 AGGAGAGAAAGATGATTAAGTCTTCCCTGTGGAAGCCATACATATGTTGGCAGGTC 548
 DB 361 AGGAGAGAAAGATGATTAAGTCTTCCCTGTGGAAGCCATACATATGTTGGCAGGTC 420
 QY 549 CTGAAAGAGATGATGATGAGGCTGAGCCAGCTGAGCTTACTCTCATCTTCTTCT 608
 DB 421 CTGAAAGAGATGATGATGAGGCTGAGCCAGCTGAGCTTACTCTCATCTTCTTCT 480
 QY 609 CATGTGACCTGTGTAAGAGCTTGAATTCAGGCTCATTTGAGGCTCACTAGTATGAGA 668
 DB 481 CATGTGACCTGTGTAAGAGCTTGAATTCAGGCTCATTTGAGGCTCACTAGTATGAGA 540
 QY 669 GAAGGAGTCTGGCCAGAGAAAGACACACCTTGCACAAATTTATCTATTTTGGT 728
 DB 541 GAAGGAGTCTGGCCAGAGAAAGACACACCTTGCACAAATTTATCTATTTTGGT 600
 QY 729 GTATTGTGAAGGAAAGTTGGCAGCTCAGAAACAAAGACTCTGATGATCAGATAGG 788
 DB 601 GTATTGTGAAGGAAAGTTGGCAGCTCAGAAACAAAGACTCTGATGATCAGATAGG 660
 QY 789 GATGCTGATCTGCTGGGCTGAGCTTAAATGACACAGTCAATGTTATGTAACAGG 848
 DB 661 GATGCTGATCTGCTGGGCTGAGCTTAAATGACACAGTCAATGTTATGTAACAGG 720

QY 849 TCTGTCCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 908
 DB 721 TCTGTCCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
 QY 909 GGCACACTCTCTGAGAGTGCATCAATATTCCTGAGAGTGCATCAATATTCCTGAGAG 968
 DB 781 GGCACACTCTCTGAGAGTGCATCAATATTCCTGAGAGTGCATCAATATTCCTGAGAG 840
 QY 969 CATGCGAGGCTCTTGAATTCGCGCAATATTCCTTCTACTGCTCAACACTTTG 1028
 DB 841 CATGCGAGGCTCTTGAATTCGCGCAATATTCCTTCTACTGCTCAACACTTTG 900
 QY 1029 ATGACCTTGAAGATTTCTACTGTTTGTGATATCTCTCCACCAATGATGAGGATG 1088
 DB 901 ATGACCTTGAAGATTTCTACTGTTTGTGATATCTCTCCACCAATGATGAGGATG 960
 QY 1089 GAAGCTTATGCAAGATGAGAGCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1148
 DB 961 GAAGCTTATGCAAGATGAGAGCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
 QY 1149 GAAGAGCGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1208
 DB 1021 GAAGAGCGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
 QY 1209 GATGATGACAACT 1268
 DB 1081 GATGATGACAACT 1140
 QY 1269 ACTTGGTATCATTAATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1328
 DB 1141 ACTTGGTATCATTAATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
 QY 1329 GCCCGGATGACAGAGTATTAAGTCAATATTTGAACAATGGCCCTCAGCGGATGGT 1388
 DB 1201 GCCCGGATGACAGAGTATTAAGTCAATATTTGAACAATGGCCCTCAGCGGATGGT 1260
 QY 1389 AGGAAGTACAAAAGAGTCCATTTATGCAATGACATGAAACCTTAAGACTCGTGA 1448
 DB 1261 AGGAAGTACAAAAGAGTCCATTTATGCAATGACATGAAACCTTAAGACTCGTGA 1320
 QY 1449 GCTATTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1508
 DB 1321 GCTATTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 QY 1509 CTGTGATTAATTTAAGAAATCAAGAGAGACCATTAATCAATCTACCGTACAGGATC 1568
 DB 1381 CTGTGATTAATTTAAGAAATCAAGAGAGACCATTAATCAATCTACCGTACAGGATC 1440
 QY 1569 ACTGATGCTCGTCTTGTGATCAAGAGATTAACAAAGGTGTAACATTTGAAGAT 1628
 DB 1441 ACTGATGCTCGTCTTGTGATCAAGAGATTAACAAAGGTGTAACATTTGAAGAT 1500
 QY 1629 TTTCAATTCGCGAGAGAGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1688
 DB 1501 TTTCAATTCGCGAGAGAGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1560
 QY 1689 CCAACTAATTCAGATCTCGTGGTCTGACCGCTATTAATCTACTGATTTGTTAATGAG 1748
 DB 1561 CCAACTAATTCAGATCTCGTGGTCTGACCGCTATTAATCTACTGATTTGTTAATGAG 1620
 QY 1749 AGAGATCTAGCTTGAAGATCTTGGCCCTCTCTCATCTGCTCAAAAGATCTGTAGAT 1808
 DB 1621 AGAGATCTAGCTTGAAGATCTTGGCCCTCTCTCATCTGCTCAAAAGATCTGTAGAT 1680
 QY 1809 CAAAGAGAGAAACGATTAATGTCAGACAAGAGAAATGATCTGTTTCTGTTTATGAT 1868
 DB 1681 CAAAGAGAGAAACGATTAATGTCAGACAAGAGAAATGATCTGTTTCTGTTTATGAT 1740
 QY 1869 GAGAACCAGAGCTGTACTCACAGAGAAATATACAGCTTCTCCCAATCCAGCTGGA 1928
 DB 1741 GAGAACCAGAGCTGTACTCACAGAGAAATATACAGCTTCTCCCAATCCAGCTGGA 1800
 QY 1929 GTGAGCTTGAAGATTCAGAGTTCAGAGCTTCCAACTCAATCATGACAGATCAATGCTAT 1988

QY 1449 GCTATCAGCATGATGAGGAATCTTGGACCTTACTTTATGCGGGAAGTTGGAGACACA 1508
| | | | |
Db 1324 GCTATTCAGATGATGATGAGAAATCTTGGACCTTACTTTATGCGGGAAGTTGGAGACACA 1383
QY 1509 CTGTGATATATATTAAGAAATCAAGACAGACATATTAACATCTACCTCAGCGGAATC 1568
| | | | |
Db 1384 CTGTGATATATATTAAGAAATCAAGACAGACATATTAACATCTACCTCAGCGGAATC 1443
QY 1569 ACATGATCTCGCTCTTGTGATTCAGAGAGATTAACAAAAGGTGTAACATTTGAAAGAT 1628
| | | | |
Db 1444 ACTGATGTCGCTCTTGTGATTCAGAGAGATTAACAAAAGGTGTAACATTTGAAAGAT 1503
QY 1629 TTTCCATTCGCGAGGAGAAATATTAATTAATGAAGACAGTACTGTAGAAGATGG 1688
| | | | |
Db 1504 TTTCCATTCGCGAGGAGAAATATTAATTAATGAAGACAGTACTGTAGAAGATGG 1563
QY 1689 CCAACTAAATCAGATCCTCGGTGCTGACCCGCTATTACTCTAGTTTCGTTAATGAGAG 1748
| | | | |
Db 1564 CCAACTAAATCAGATCCTCGGTGCTGACCCGCTATTACTCTAGTTTCGTTAATGAGAG 1623
QY 1749 AGGATCTACCTTCAGAGACTCATTTGGCCCTCTCTCATCTGTACAAAGAAATCTGAGAT 1808
| | | | |
Db 1624 AGGATCTACCTTCAGAGACTCATTTGGCCCTCTCTCATCTGTACAAAGAAATCTGAGAT 1683
QY 1809 CAAGAGGAAACAGATTAATGTGACAGACAGAAATGTCACTCTGTTTCTGATTTGAT 1868
| | | | |
Db 1684 CAAGAGGAAACAGATTAATGTGACAGACAGAAATGTCACTCTGTTTCTGATTTGAT 1743
QY 1869 GAGAACCGAAGCTGTACTCTCACAGAAATATACAGCGTTTCTCCCAATCAGCTGGA 1928
| | | | |
Db 1744 GAGAACCGAAGCTGTACTCTCACAGAAATATACAGCGTTTCTCCCAATCAGCTGGA 1803
QY 1929 GTGCACCTTGAGAGATCCAGAGTTCCAAAGCTCCAAATCATGCAAGCATCAATGGCTAT 1988
| | | | |
Db 1804 GTGCACCTTGAGAGATCCAGAGTTCCAAAGCTCCAAATCATGCAAGCATCAATGGCTAT 1863
QY 1989 GTTTTGTAGTTGTCAGTTGTGAGTTGTTGTCATGAGGTGTCACAGATCAATTTCA 2048
| | | | |
Db 1864 GTTTTGTAGTTGTCAGTTGTGAGTTGTTGTCATGAGGTGTCACAGATCAATTTCA 1923
QY 2049 AGCATTTGAGACACAGACTACTCTTCTGTCCTTCTCTGAGATATACCTTCAAAAC 2108
| | | | |
Db 1924 AGCATTTGAGACACAGACTACTCTTCTGTCCTTCTCTGAGATATACCTTCAAAAC 1983
QY 2109 AAAATGCTATGAGACACACACTACCTTATTCCTTCTCTGAGAGAACTGTTCTATG 2168
| | | | |
Db 1984 AAAATGCTATGAGACACACACTACCTTATTCCTTCTCTGAGAGAACTGTTCTATG 2043
QY 2169 TCGATGGAACCCAGAGTCTATGATTTGCGGTGTCACAACTCAGACTTTGGGAACAGA 2228
| | | | |
Db 2044 TCGATGGAACCCAGAGTCTATGATTTGCGGTGTCACAACTCAGACTTTGGGAACAGA 2103
QY 2229 GGCATGACCCGCTTACTGAGGTTTCTAGTTGTGACAAAGACACTGTGATATTACGAG 2288
| | | | |
Db 2104 GGCATGACCCGCTTACTGAGGTTTCTAGTTGTGACAAAGACACTGTGATATTACGAG 2163
QY 2289 GACATTTATGAGATATTTACGATATCTGCTGAGTAAAAAAATGCTTGAACCAAGA 2348
| | | | |
Db 2164 GACATTTATGAGATATTTACGATATCTGCTGAGTAAAAAAATGCTTGAACCAAGA 2223
QY 2349 AGCTTCTCCAGAACCCACAGTCTTGAAGAGCCATCAACGGGAATATACGTGACTACT 2408
| | | | |
Db 2224 -----GAAATTAACGTGCTACT 2241
QY 2409 CTTCAGTCAAGATCAAGAGAAATTAATGATGATACCATATCAAGTTGAATGAAGAG 2468
| | | | |
Db 2242 CTTCAGTCAAGATCAAGAGAAATTAATGATGATACCATATCAAGTTGAATGAAGAG 2301
QY 2469 GAAGATTTGACATTTATGATGAGATGAAAAATGAGAGCCCCCGAGCTTTCAAAAGAAA 2528
| | | | |
Db 2302 GAAGATTTGACATTTATGATGAGATGAAAAATGAGAGCCCCCGAGCTTTCAAAAGAAA 2361

QY 2529 ACACGACACTATTTTATTCCTGACATGAGAGGCTCTGGGATTTATGGATGATGAGTCTC 2588
| | | | |
Db 2362 ACACGACACTATTTTATTCCTGACATGAGAGGCTCTGGGATTTATGGATGATGAGTCTC 2421
QY 2589 CCACATGCTCTAAGAAACAGGCGCTCAGAGTGGCATGTCTCCATCTCAAGAAATGTT 2648
| | | | |
Db 2422 CCACATGCTCTAAGAAACAGGCGCTCAGAGTGGCATGTCTCCATCTCAAGAAATGTT 2481
QY 2649 TTCAGAGAAATTTACGATGAGTCTCTTACTAGCCCTTATACCCGTGGAAGAACTAAATGAA 2708
| | | | |
Db 2482 TTCAGAGAAATTTACGATGAGTCTCTTACTAGCCCTTATACCCGTGGAAGAACTAAATGAA 2541
QY 2709 CATTTGGGACCTCTGGGGCCATATATTAAGACAGAGAGTTGAAGATATATCATGTAAT 2768
| | | | |
Db 2542 CATTTGGGACCTCTGGGGCCATATATTAAGACAGAGAGTTGAAGATATATCATGTAAT 2601
QY 2769 TTCAGAAATCAGGCGCTCTGCTGCTATTCCTTCTATTTCTAGCCTTATTTCTTATGAGGAA 2828
| | | | |
Db 2602 TTCAGAAATCAGGCGCTCTGCTGCTATTCCTTCTATTTCTAGCCTTATTTCTTATGAGGAA 2661
QY 2829 GATCAGAGGCAAGGAGCAGAACCTTGAAGAAACCTTGTCAAGCCTAATGAAGAACCAACT 2888
| | | | |
Db 2662 GATCAGAGGCAAGGAGCAGAACCTTGAAGAAACCTTGTCAAGCCTAATGAAGAACCAACT 2721
QY 2889 TACTTTTGAAGATGCAACATCATATGACACCCCACTAAAGATGATTTGACTGCAAAAGCC 2948
| | | | |
Db 2722 TACTTTTGAAGATGCAACATCATATGACACCCCACTAAAGATGATTTGACTGCAAAAGCC 2781
QY 2949 TGGGCTTATTTCTGATGTTGACTGGAAGAAAGATGTCACCTCAGGCTTATTTGGAGCC 3008
| | | | |
Db 2782 TGGGCTTATTTCTGATGTTGACTGGAAGAAAGATGTCACCTCAGGCTTATTTGGAGCC 2841
QY 3009 CTTCGTGCTGACACACTATACACTGAACCTGCTCATGAGGAGACAAAGTACAGTACAG 3068
| | | | |
Db 2842 CTTCGTGCTGACACACTATACACTGAACCTGCTCATGAGGAGACAAAGTACAGTACAG 2901
QY 3069 GAATTTGCTGTTTATACCATCTTTGATGAGACAAAGCTGTAACCTGACAGAAAT 3128
| | | | |
Db 2902 GAATTTGCTGTTTCTGATGTTGACTGGAAGAAAGTGTGCACTAGGCTTATTTGGAGCC 2961
QY 3129 ATGGAAGAAACTGCAAGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTAAAGAG 3188
| | | | |
Db 2962 ATGGAAGAAACTGCAAGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTAAAGAG 3021
QY 3189 AATTAATGCTTCCATGCAATCAATGCTTACATTAATGATACACTAGCTGCTTATG 3248
| | | | |
Db 3022 AATTAATGCTTCCATGCAATCAATGCTTACATTAATGATACACTAGCTGCTTATG 3081
QY 3249 GCTCAGATCAAAAGATTCAGTGTATCTGCTCAGCATGGGACAGCAATGAAGACATCCAT 3308
| | | | |
Db 3082 GCTCAGATCAAAAGATTCAGTGTATCTGCTCAGCATGGGACAGCAATGAAGACATCCAT 3141
QY 3309 TCTATTCATTTCACTGAGCATGTTGCTACGTATGAGAAAGAGGAGTAAATATGGA 3368
| | | | |
Db 3142 TCTATTCATTTCACTGAGCATGTTGCTACGTATGAGAAAGAGGAGTAAATATGGA 3201
QY 3369 CTGTACATCTCTATCCAGGTGTTTGAAGACAGTGAAGATGTTACCATCCAAAGCTGGA 3428
| | | | |
Db 3202 CTGTACATCTCTATCCAGGTGTTTGAAGACAGTGAAGATGTTACCATCCAAAGCTGGA 3261
QY 3429 AATTTGGGGGTGGGAATGCTTATTTGGGAGCATCTACATGCTGGGATGAGACACTTTT 3488
| | | | |
Db 3262 AATTTGGGGGTGGGAATGCTTATTTGGGAGCATCTACATGCTGGGATGAGACACTTTT 3321
QY 3489 CTGTGTACAGCAATTAAGTGTACAGATCCCTGCGGAATGGCTCTGGAACATTAGAGAT 3548
| | | | |
Db 3322 CTGTGTACAGCAATTAAGTGTACAGATCCCTGCGGAATGGCTCTGGAACATTAGAGAT 3381
QY 3549 TTTCAAGATTAACAGCTTACAGACAAATATGACAGTGGGCCCCCAAGCTGGCCACACTTCAT 3608
| | | | |
Db 3382 TTTCAAGATTAACAGCTTACAGACAAATATGACAGTGGGCCCCCAAGCTGGCCACACTTCAT 3441
QY 3609 TATTCGGGATCAATCAATGCTGTGAGACCAAGAGGCCCTTTCTTGATCAAGGTGAT 3668

```

Db 3442 TATTCGGATCATTAATGCGCTGGAGACCAAGAGCCCTTTTCTTGATCAGATGAT 3501
Oy 3669 CTGTGGACCAATGATTATTCACGGCATCAAGACCAAGGGTCCCGTCAGAAATTC 3728
Db 3502 CTGTGGACCAATGATTATTCACGGCATCAAGACCAAGGGTCCCGTCAGAAATTC 3561
Oy 3729 AGCCTCTACATCTCTCAATTTATCATCATGTTATCTTGATGGAGAAATGGCAGACT 3788
Db 3562 AGCCTCTACATCTCTCAATTTATCATCATGTTATCTTGATGGAGAAATGGCAGACT 3621
Oy 3789 TATCGAGAAATTCACATGGAACCTTAATGGCTCTTGAGCAATGGATTCATCTGG 3848
Db 3662 TATCGAGAAATTCACATGGAACCTTAATGGCTCTTGAGCAATGGATTCATCTGG 3681
Oy 3849 ATAAACACAATATTTTAAACCTCAATTAATGCTGATACATCCGTTGCAACCAACT 3908
Db 3682 ATAAACACAATATTTTAAACCTCAATTAATGCTGATACATCCGTTGCAACCAACT 3741
Oy 3909 CATTTACCATTCGACAGACTCTTGATGGAGTTGATGGGTGATTTAAATAGTTGC 3968
Db 3742 CATTTACCATTCGACAGACTCTTGATGGAGTTGATGGGTGATTTAAATAGTTGC 3801
Oy 3969 AGCATGCGATGGAGATGAGAGTAAGCAATATCAGATGCACAGATTACTGCTCATCC 4028
Db 3802 AGCATGCGATGGAGATGAGAGTAAGCAATATCAGATGCACAGATTACTGCTCATCC 3861
Oy 4029 TACTTTACCAATATGTTTGCACCTGCTGCTTCAAAAAGCTGCACTCACCTCCAAAGG 4088
Db 3862 TACTTTACCAATATGTTTGCACCTGCTGCTTCAAAAAGCTGCACTCACCTCCAAAGG 3921
Oy 4089 AGCATATGCTGAGAGACTGAGTGAATATCCAAAAGAGTGGCTGCAATGGACTTC 4148
Db 3922 AGGATATGCTGAGAGACTGAGTGAATATCCAAAAGAGTGGCTGCAATGGACTTC 3981
Oy 4149 CAGAAGACAATGAAGTACAGAGAGTAACTACTAGGAGTAATCTCTGTTACCAAGC 4208
Db 3982 CAGAAGACAATGAAGTACAGAGAGTAACTACTAGGAGTAATCTCTGTTACCAAGC 4041
Oy 4209 ATGATATGAAGAGTCTCTCATCTCCAGACATCAAGATGGCCATCAGTGCATCTCTT 4268
Db 4042 ATGATATGAAGAGTCTCTCATCTCCAGACATCAAGATGGCCATCAGTGCATCTCTT 4101
Oy 4269 TTTGGAATGGCAAGTAAAGTCTTTTCAAGGAAATCAAGACTCCTTCAACCTGTGGT 4328
Db 4102 TTTGGAATGGCAAGTAAAGTCTTTTCAAGGAAATCAAGACTCCTTCAACCTGTGGT 4161
Oy 4329 AACCTCTAGACCCACCGTACTGACTGCTGCTGCAATCAGCCCAAGGTTGGGTG 4388
Db 4162 AACCTCTAGACCCACCGTACTGACTGCTGCTGCAATCAGCCCAAGGTTGGGTG 4221
Oy 4389 CACCAAGATGGCCCTGAGATGGAGTTCTGGGCTCGAGGCAAGACCTCTAC 4442
Db 4222 CACCAAGATGGCCCTGAGATGGAGTTCTGGGCTCGAGGCAAGACCTCTAC 4275

```

RESULT 9

```

N80447 standard; DNA: 4272 BP.
N80447:
10-07-1990 (first entry)
DE Modified factor VIII:C sequence with the R740-D1658 deletion.
RM Modified factor VIII:C; haemophilia; procoagulant;
OS blood coagulation; RD deletion; ss.
OS Homo sapiens.
PM M0800831-A.
FU 11-FEB-1988.
FR 31-JUL-1987; U01814.
ER 01-AUG-1986; US-893375.
PA (BioJ) Blegen NV (PASE/).
NT PASEK MP.
DR MPI: 88-049866/07.
P-PsDB: P80266.

```

```

PT New DNA sequences encoding modified factor VIII:C - with deletion of DNA
PT encoding maturation polypeptide, useful for high yield transformation.
PS Claim 3; Page 47-48-49-50; 97pp; English.
PS The RD deletion removes the DNA from Ser 741 to Ser 1657. A major part
CC of the sequence encoding the maturation polypeptide of
CC factor VIII:C is deleted, i.e. Gln 744 - Asp 1653.
CC The full length Factor VIII:C cDNA has two changes with respect to the
CC published sequence (EPO application 160457):
CC CTG to CTA at Leu 242 and TTC to CTC change at amino acid residue 1880
CC (Phe to Leu). The product is produced in approx. 20 times higher
CC yields than previous recombinant produced factor VIII:C and are more
CC easily purified. The peptide is used for treating haemophilia A, both
CC acute and prolonged bleeding.
CC See also N80444 and N80446.
SQ Sequence 4272 BP; 1243 A; 941 C; 944 G; 1144 T;

Query Match      87.2%; Score 4213.6; DB 1; Length 4272;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 4268; Conservative 0; Mismatches 4; Indels 42; Gaps 1;

Oy 129 GCCACCAAGAAATACCTAGCTGGTGCAGTGAATGTCATGGAGTATATGCAAGTAT 188
Db 1 GCCACCAAGAAATACCTAGCTGGTGCAGTGAATGTCATGGAGTATATGCAAGTAT 60
Oy 189 CTCGGTGAAGTCTGCTGAGCGCAAGATTCTCTTAAGTGCCTTCTTCCATTC 248
Db 61 CTCGGTGAAGTCTGCTGAGCGCAAGATTCTCTTAAGTGCCTTCTTCCATTC 120
Oy 249 AACACCTAGTGTGTACAAAAGACCTGTTTGTAGAAATCAGGATCACCCTTTCAC 308
Db 121 AACACCTAGTGTGTACAAAAGACCTGTTTGTAGAAATCAGGATCACCCTTTCAC 180
Oy 309 ATCGCTAAGCCAAAGCCACCTGATGGGTGCTGTAGTCTCTTACCATCAGGCTGAGTT 368
Db 181 ATCGCTAAGCCAAAGCCACCTGATGGGTGCTGTAGTCTCTTACCATCAGGCTGAGTT 240
Oy 369 TATGATACAGTGTGTATTAACCTTAAGAAATGCTTCCATCTGTCATGCTTATGCT 428
Db 241 TATGATACAGTGTGTATTAACCTTAAGAAATGCTTCCATCTGTCATGCTTATGCT 300
Oy 429 GTTGTGTATCTACTGAAAGCTCTTGAAGAGTGTGAATATGATGATCAGCAAGTCA 488
Db 301 GTTGTGTATCTACTGAAAGCTCTTGAAGAGTGTGAATATGATGATCAGCAAGTCA 360
Oy 489 AGGGAAGAAAGATGATTAAGTCTTCCCTGTGGAACCAATCATATGTCGCGAGTGC 548
Db 361 AGGGAAGAAAGATGATTAAGTCTTCCCTGTGGAACCAATCATATGTCGCGAGTGC 420
Oy 549 CTGAAAGAAATGTCATAGGCTCTGACCCACTGTGCTTACCTACTCATATCTTCT 608
Db 421 CTGAAAGAAATGTCATAGGCTCTGACCCACTGTGCTTACCTACTCATATCTTCT 480
Oy 609 CATGTGACCTGTGTAAGACTGTAATCAGGCCCTTGAAGCCCTACTATAGTGA 668
Db 481 CATGTGACCTGTGTAAGACTGTAATCAGGCCCTTGAAGCCCTACTATAGTGA 540
Oy 669 GAAGGAGCTGTGCGCAAGAAAGACACAGACTTGCACAAATTTATCTACTTTTCT 728
Db 541 GAAGGAGCTGTGCGCAAGAAAGACACAGACTTGCACAAATTTATCTACTTTTCT 600
Oy 729 GTATTGTATGAAGGAAAGTTGGCAGTCAAGAACTCTTATGATGAGATAGG 788
Db 601 GTATTGTATGAAGGAAAGTTGGCAGTCAAGAACTCTTATGATGAGATAGG 660
Oy 789 GATGCTGATCTGCTCGGGCTCTGGCTTAAATGACACAGTCAATGTTATGTAACAG 848
Db 661 GATGCTGATCTGCTCGGGCTCTGGCTTAAATGACACAGTCAATGTTATGTAACAG 720
Oy 849 TCTCTGCAAGTCTGATTTGATGACCAAGAAATCAGTCTATTTGATGATGATGAT 908
Db 721 TCTCTGCAAGTCTGATTTGATGACCAAGAAATCAGTCTATTTGATGATGATGAT 780

```


QY 909 GGCACACTCTGTAAGTGCATCAATATTCCTGCAAGTGCACACATTTCTTGAGGAAAC 968
| | | | |
Db 781 GGCACACTCTGTAAGTGCATCAATATTCCTGCAAGTGCACACATTTCTTGAGGAAAC 840
| | | | |
QY 969 CATTGCGAGCGTCTTGGAATCTCGCAATTAATTTCTTACTGTCGAACGCTCTTG 1028
| | | | |
Db 841 CATGCGAGCGTCTTGGAATCTCGCAATTAATTTCTTACTGTCGAACGCTCTTG 900
| | | | |
QY 1029 ATGACCTTGACAGTTTCTACTGTTTGTATATCTCTTCCACCAATGATGCGATG 1088
| | | | |
Db 901 ATGACCTTGACAGTTTCTACTGTTTGTATATCTCTTCCACCAATGATGCGATG 960
| | | | |
QY 1089 GAACCTATGCAAGTAGACAGTGTCCAGAGAACCCCACTACCAATGAATAAT 1148
| | | | |
Db 961 GAACCTATGCAAGTAGACAGTGTCCAGAGAACCCCACTACCAATGAATAAT 1020
| | | | |
QY 1149 GAAGAACGGAAGACATATGATGATCTTACTGATCTGAAATGATGTCAGGTTT 1208
| | | | |
1021 GAAGAACGGAAGACATATGATGATCTTACTGATCTGAAATGATGTCAGGTTT 1080
| | | | |
QY 1209 GATGATACAACTCTCTCTCTTATCCAAATTCGTCAGTTGCGAAGAACATCTTAA 1268
| | | | |
Db 1081 GATGATACAACTCTCTCTCTTATCCAAATTCGTCAGTTGCGAAGAACATCTTAA 1140
| | | | |
QY 1269 ACTTGGTACATTACATTTGCTGCGAGAGAGAGACTGGGACTATGCTCCCTTACTC 1328
| | | | |
Db 1141 ACTTGGTACATTACATTTGCTGCGAGAGAGAGACTGGGACTATGCTCCCTTACTC 1200
| | | | |
QY 1329 GCCCCGATGACAGAAAGTTATTAAGTCATATTTGAAACATGCGCTCAGCGATGGT 1388
| | | | |
Db 1201 GCCCCGATGACAGAAAGTTATTAAGTCATATTTGAAACATGCGCTCAGCGATGGT 1260
| | | | |
QY 1389 AGGAAGTACAAAAAGTCCCATTTATGCGATACACAGATGAAACCTTTAAGACTGTA 1448
| | | | |
Db 1261 AGGAAGTACAAAAAGTCCCATTTATGCGATACACAGATGAAACCTTTAAGACTGTA 1320
| | | | |
QY 1449 GCATTCAGATGAATCAGGAATCTTGCGACCTTACTTATGCGGAGTTGAGACACA 1508
| | | | |
Db 1321 GCATTCAGATGAATCAGGAATCTTGCGACCTTACTTATGCGGAGTTGAGACACA 1380
| | | | |
QY 1509 CTGTGATTAATATTAAGAAATCAAGACAGACATATTAACATCTTACCTCAGGAATC 1568
| | | | |
Db 1381 CTGTGATTAATATTAAGAAATCAAGACAGACATATTAACATCTTACCTCAGGAATC 1440
| | | | |
QY 1569 ACAGAGTCCGCTCTGTTGTTTCAAGAGATTTCCAAAAGCTTAAACATTTGAAGAT 1628
| | | | |
1441 ACAGAGTCCGCTCTGTTGTTTCAAGAGATTTCCAAAAGCTTAAACATTTGAAGAT 1500
| | | | |
QY 1629 TTTCAATTCGCGAGAGAAATATTCAAATTAATGACAGTGACTGAGAGAGATGG 1688
| | | | |
Db 1501 TTTCAATTCGCGAGAGAAATATTCAAATTAATGACAGTGACTGAGAGATGG 1560
| | | | |
QY 1689 CCAACTTAATCAGATCCTCGGTGCTGACCCGCTATTAATCTAGTTTCGTTAATGAG 1748
| | | | |
Db 1561 CCAACTTAATCAGATCCTCGGTGCTGACCCGCTATTAATCTAGTTTCGTTAATGAG 1620
| | | | |
QY 1749 AGAGATCTGCTCAGAGCTCATTTGGCCCTCTCTATCTGCTCAAGAAATCTGAT 1808
| | | | |
Db 1621 AGAGATCTGCTCAGAGCTCATTTGGCCCTCTCTATCTGCTCAAGAAATCTGAT 1680
| | | | |
QY 1809 CAAAGAGAAACAGATTAATGTGACAGACAGAGAGATGTATCTGTTTCTGATTTGAT 1868
| | | | |
Db 1681 CAAAGAGAAACAGATTAATGTGACAGACAGAGAGATGTATCTGTTTCTGATTTGAT 1740
| | | | |
QY 1869 GAGAACCGAAGCTGTACTCACAGAGAAATATCAACGCTTCTCCCAATCCAGCTGA 1928
| | | | |
1741 GAGAACCGAAGCTGTACTCACAGAGAAATATCAACGCTTCTCCCAATCCAGCTGA 1800
| | | | |
QY 1929 GTGACGCTGAGAGTCCAGAGTTCCAGGCTCCCAATCATGACAGCATCAATGCTAT 1988
| | | | |
Db 1801 GTGACGCTGAGAGTCCAGAGTTCCAGGCTCCCAATCATGACAGCATCAATGCTAT 1860
| | | | |
QY 1989 GTTTTGTAGTTTGCAGTTGTGATGATGAGGTGACATACATGATCAATCA 2048
| | | | |

Db 1861 GTTTTGTAGTTTGCAGTTGTCAGTTGTTGATGAGAGTGGACATCAATGATCA 1920
| | | | |
QY 2049 AGCATGGAGCAGACACTACTTCTTCTGTCTTCTTCTGATATACCTTCAACAC 2108
| | | | |
Db 1921 AGCATGGAGCAGACACTACTTCTTCTGTCTTCTTCTGATATACCTTCAACAC 1980
| | | | |
QY 2109 AAAATGCTATGAAGACACACTACCTATTCCTCAGGAGAAACCTGCTCATG 2168
| | | | |
Db 1981 AAAATGCTATGAAGACACACTACCTATTCCTCAGGAGAAACCTGCTCATG 2040
| | | | |
QY 2169 TCGATGAAAACCCAGCTGTATGATTTGCGGTGCGCAACTCAGACTTTGGAGACA 2228
| | | | |
Db 2041 TCGATGAAAACCCAGCTGTATGATTTGCGGTGCGCAACTCAGACTTTGGAGACA 2100
| | | | |
QY 2229 GGCATACCGCTTACTGAGGTTTCTAGTTGACAAACACACGCTGATTTACGAG 2288
| | | | |
Db 2101 GGCATACCGCTTACTGAGGTTTCTAGTTGACAAACACGCTGATTTACGAG 2160
| | | | |
QY 2289 GACAGTTATGAAGATATTTACAGCATCTGCTGAGTAAACAAATGCCATTGAACAGA 2348
| | | | |
Db 2161 GACAGTTATGAAGATATTTACAGCATCTGCTGAGTAAACAAATGCCATTGAACAGA 2220
| | | | |
QY 2349 AGCTTCTCCAGAACCCAGCTCTTGAACGCCATCAACGGAAATTAACGTACTACT 2408
| | | | |
2221 -----GAAATTAACGTACTACT 2238
| | | | |
QY 2409 CTTGAGTCAGATCAAGAGAAATGACTATGATGATACCATTCAGTTGAATGAAGA 2468
| | | | |
Db 2239 CTTGAGTCAGATCAAGAGAAATGACTATGATGATACCATTCAGTTGAATGAAGA 2298
| | | | |
QY 2469 GAAGATTTTACATTTATGATGAGATGAAATTCAGAGCCCGCAGCTTCAAAAAGAA 2528
| | | | |
Db 2299 GAAGATTTTACATTTATGATGAGATGAAATTCAGAGCCCGCAGCTTCAAAAAGAA 2358
| | | | |
QY 2529 ACACGACATATTTATTTATGCTCAGATGAGAGGCTCTGGGATTAAGGATGATGCTC 2588
| | | | |
Db 2359 ACACGACATATTTATTTATGCTCAGATGAGAGGCTCTGGGATTAAGGATGATGCTC 2418
| | | | |
QY 2589 CCACATGTTCTAAGAAACAGGCTCAGAGTGGGACTGTCCTCAATTAAGAAATGTT 2648
| | | | |
Db 2419 CCACATGTTCTAAGAAACAGGCTCAGAGTGGGACTGTCCTCAATTAAGAAATGTT 2478
| | | | |
QY 2649 TTCCAGGAATTTAGTATGCTCTTACTACAGCCCTTATACCGTGGAGAACTAATGA 2708
| | | | |
Db 2479 TTCCAGGAATTTAGTATGCTCTTACTACAGCCCTTATACCGTGGAGAACTAATGA 2538
| | | | |
QY 2709 CATTTGGAGCTCTGCGGCATATTAAGAGCAGAAAGTTGAAGATATATCATGCTA 2768
| | | | |
Db 2539 CATTTGGAGCTCTGCGGCATATTAAGAGCAGAAAGTTGAAGATATATCATGCTA 2598
| | | | |
QY 2769 TTCAGAAATCAGGCTCTGTCGCCATATTAAGAGCAGAAAGTTGAAGATATATCAT 2828
| | | | |
Db 2599 TTCAGAAATCAGGCTCTGTCGCCATATTAAGAGCAGAAAGTTGAAGATATATCAT 2658
| | | | |
QY 2829 GATCAGAGCAGAGCAGAACTAGAAAAAACTTTGCAAGCCTATATGAACCAAACT 2888
| | | | |
Db 2659 GATCAGAGCAGAGCAGAACTAGAAAAAACTTTGCAAGCCTATATGAACCAAACT 2718
| | | | |
QY 2889 TACTTTTGGAAAGTCAACATATATGACACCCACTAAAGATGATTTGATGCAAAAGC 2948
| | | | |
Db 2719 TACTTTTGGAAAGTCAACATATATGACACCCACTAAAGATGATTTGATGCAAAAGC 2778
| | | | |
QY 2949 TGGGCTATTTCTGATGTTGACCTGGAAGAAAGATGCACTCAGGCTGATTTGAGACC 3008
| | | | |
Db 2779 TGGGCTATTTCTGATGTTGACCTGGAAGAAAGATGCACTCAGGCTGATTTGAGACC 2838
| | | | |
QY 3009 CTTCTGCTGCCACACTAACAACCTGCTCATGAGAGACAAAGTACAGATACAG 3068
| | | | |
Db 2839 CTTCTGCTGCCACACTAACAACCTGCTCATGAGAGACAAAGTACAGATACAG 2898
| | | | |
QY 3069 GAATTTGCTGTTTTCACCATCTTTGATGAGACCAAAAGCTGTACTTCACTGAAT 3128
| | | | |

Db 2899 GAATTTGCTGTTTTTACCACCTTTTGATGAGACAAAGGCTGTAATCTACTGAAAT 2958
 QY 3129 ATGGAAGAAAGAACTGAGGCGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAG 3188
 Db 2959 ATGGAAGAAAGAACTGAGGCGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAG 3018
 QY 3189 AATATGCTGCTCCATGCAATCAATGCTTACATATGATGATCACTGCTGTTATG 3248
 Db 3019 AATTATGCTTCCATGCAATCAATGCTTACATATGATGATCACTGCTGTTATG 3078
 QY 3249 GCTCAGAGTCAAAAGAGATTCGATGCTGCTGAGCATGGGCGACCAATGAAACATCAT 3308
 Db 3079 GCTCAGAGTCAAAAGAGATTCGATGCTGCTGAGCATGGGCGACCAATGAAACATCAT 3138
 QY 3309 TCTATTATTTCAGTGAACATGTTGTTCACTGTACGAAAAAAGAGATATAATGGCA 3368
 Db 3139 TCTATTATTTCAGTGAACATGTTGTTCACTGTACGAAAAAAGAGATATAATGGCA 3198
 QY 3369 CTGTACACATCTATCCAGGCTTTTGAGACAGTGGAAATGTACATCCAAAGCTGGA 3428
 Db 3199 CTGTACACATCTATCCAGGCTTTTGAGACAGTGGAAATGTACATCCAAAGCTGGA 3258
 QY 3429 ATTTGGCGGTGGAATGCTTATTTGCGAGCATCTACATGCTGGGATGAGACACTTTT 3488
 Db 3259 ATTTGGCGGTGGAATGCTTATTTGCGAGCATCTACATGCTGGGATGAGACACTTTT 3318
 QY 3489 CTGTGTACAGCAATTAAGTGTACAGTCTCCCTGGGAATGCTTGTGACACATTAGAGAT 3548
 Db 3319 CTGTGTGTACAGCAATTAAGTGTACAGTCTCCCTGGGAATGCTTGTGACACATTAGAGAT 3378
 QY 3549 TTTTACAGTTACAGTTGAGCAATATGAGACAGTGGGCCCCAAAGCTGGCGACACTTCAT 3608
 Db 3379 TTTTACAGTTACAGTTGAGCAATATGAGACAGTGGGCCCCAAAGCTGGCGACACTTCAT 3438
 QY 3609 TATTCGCGATCAATCAATGCTGTGAGACCAAGAGGCTTTTCTTGATCAAGTGGAT 3668
 Db 3439 TATTCGCGATCAATCAATGCTGTGAGACCAAGAGGCTTTTCTTGATCAAGTGGAT 3498
 QY 3669 CTGTGGGACCAATGATTTATTCACGGCATCAAGACCCAGGGTGGCCCTGAGAGTTCTC 3728
 Db 3499 CTGTGGGACCAATGATTTATTCACGGCATCAAGACCCAGGGTGGCCCTGAGAGTTCTC 3558
 QY 3729 AGCCTTACATCTCTCAGTTTATCATGATGATGCTGTGATGGAAGAAGTGGCAGACT 3788
 Db 3559 AGCCTTACATCTCTCAGTTTATCATGATGATGCTGTGATGGAAGAAGTGGCAGACT 3618
 QY 3789 TATCGAGAAATTCACCTGGAACCTTAATGCTTCTTTGGCAATGAGATTCATCTGGG 3848
 Db 3619 TATCGAGAAATTCACCTGGAACCTTAATGCTTCTTTGGCAATGAGATTCATCTGGG 3678
 QY 3849 ATAAACACAAATATTTTAAACCTTCACCAATTAATGCTGCAATCCGTTTGCACCAACT 3908
 Db 3679 ATAAACACAAATATTTTAAACCTTCACCAATTAATGCTGCAATCCGTTTGCACCAACT 3738
 QY 3909 CATATATGCAATCCGACACACTCTCCGATGAGATGAGGCTGTGATTAATAGTTGC 3968
 Db 3739 CATATATGCAATCCGACACACTCTCCGATGAGATGAGGCTGTGATTAATAGTTGC 3798
 QY 3969 AGCATGCAATGGGAATGGAAGTAAAGCAATATCAGATGACAGATTAAGTCAATCC 4028
 Db 3799 AGCATGCAATGGGAATGGAAGTAAAGCAATATCAGATGACAGATTAAGTCAATCC 3858
 QY 4029 TACTTTACCAATATGTTTGCACCTGCTCTCTTCAAAAGCTGCACTCACCTCAAGGG 4088
 Db 3859 TACTTTACCAATATGTTTGCACCTGCTCTCTTCAAAAGCTGCACTCACCTCAAGGG 3918
 QY 4089 AGGAGTATGCTGGAAGCTCAGAGTGAATATCCAAAGAGTGGCTCAGTGGACTTC 4148
 Db 3919 AGGAGTATGCTGGAAGCTCAGAGTGAATATCCAAAGAGTGGCTCAGTGGACTTC 3978
 QY 4149 CAGAGACAAATGAAGTACAGAGATTAATCTAGAGGAGTAAATCTCTCTTACAGC 4208
 Db 3979 CAGAGACAAATGAAGTACAGAGATTAATCTAGAGGAGTAAATCTCTCTTACAGC 4038

QY 4209 ATGATGTGAAGAGTTCCTCATCTCCAGACATCAAGATGGCCATGAGTACTCTTT 4268
 Db 4039 ATGATGTGAAGAGTTCCTCATCTCCAGACATCAAGATGGCCATGAGTACTCTTT 4098
 QY 4269 TTTGAGAAATGGCAAGATGAAGGTTTTTTCAGGGAATCAAGACTCTTACACCGTGGTG 4328
 Db 4099 TTTGAGAAATGGCAAGATGAAGGTTTTTTCAGGGAATCAAGACTCTTACACCGTGGTG 4158
 QY 4329 AACTCTAGACCCACCGTTACTGACTGCTTACTGCTTGAATTCACCCAGAGTTGGTG 4388
 Db 4159 AACTCTAGACCCACCGTTACTGACTGCTTACTGCTTGAATTCACCCAGAGTTGGTG 4218
 QY 4389 CACCAGATTGCTCCCTGAGAGTGAAGTCTTGCTGCTGCGAGGACACAGACTCTAC 4442
 Db 4219 CACCAGATTGCTCCCTGAGAGTGAAGTCTTGCTGCTGCGAGGACACAGACTCTAC 4272

RESULT 10
 N80444 standard; DNA; 4545 BP.
 ID N80444;
 AC N80444;
 DT 10-OCT-1990 (first entry)
 DE Modified factor VIII:C sequence with the 0744-D1563 deletion.
 KM Modified factor VIII:C; haemophilia; procoagulant;
 KW blood coagulation; QD deletion; ss.
 OS Homo sapiens.
 PN M08000831-A.
 PD 11-FEB-1988.
 PE 31-JUL-1987; U01814.
 PR 01-AUG-1986; US-892375.
 PA (BIOJ) Biogen NV (PASE/).
 PI PASEK MP;
 DR NPI: 88-049866/07.
 DP P-PSDB: P80265.
 PT New DNA sequences encoding modified factor VIII:C - with deletion of DNA
 encoding maturation polypeptide, useful for high yield transformation.
 PS Claim 3; Page 38-39-40-41; 97pp; English.
 CC A major part of the sequence encoding the maturation polypeptide of
 CC factor VIII:C is deleted, i.e. Gln 744 - Asp 1563. The QD deletion
 CC retains approximately 90 amino acids of the maturation polypeptide
 CC (four amino acids at the N-terminal end and 86 amino acids at
 CC the C-terminal end). The full length factor VIII:C cDNA has two
 CC changes with respect to the published sequence (EPO application 160457):
 CC C to T at Leu 242 and T to C at Cys change at amino acid residue 1880
 CC (Phe to Leu). The product is produced in approx. 20 times higher
 CC yields than previous recombinant produced factor VIII:C and are more
 CC easily purified. The peptide is used for treating haemophilia A, both
 CC acute and prolonged bleeding.
 CC See also N80446 and N80447.
 SQ Sequence 4545 BP; 1345 A; 1002 C; 1003 G; 1195 T;

Query Match 84.2%; Score 4069.6; DB 1; Length 4545;
 Best Local Similarity 94.9%; Pred. No. 0;
 Matches 4310; Conservative 0; Mismatches 4; Indels 228; Gaps 1;

QY 129 GCCACGAGAAATGACTACTGCTGGGTGAGTGGAACTGTCATGAGGACTATATGCAAGTAT 188
 Db 4 GCCACGAGAAATGACTACTGCTGGGTGAGTGGAACTGTCATGAGGACTATATGCAAGTAT 63
 QY 189 CTGCGTAGCTGCTGTGAGCGCAAGATTTCTCTAGAGTGGCCAAATCTTTTCCATTC 248
 Db 64 CTGCGTAGCTGCTGTGAGCGCAAGATTTCTCTAGAGTGGCCAAATCTTTTCCATTC 123
 QY 249 AACACCTCAGTCTGTGCAAAAAGACTCTGTTTATAATTACAGGATCACTTTTAAAC 308
 Db 124 AACACCTCAGTCTGTGCAAAAAGACTCTGTTTATAATTACAGGATCACTTTTAAAC 183
 QY 309 ATCGTAAGCCAAAGGCCACCTGGATGGTCTGCTAGTCTTACATCCAGGCTGAGGTT 368
 Db 184 ATCGTAAGCCAAAGGCCACCTGGATGGTCTGCTAGTCTTACATCCAGGCTGAGGTT 243

QY 369 TATGATACAGTGGTCACTTACACTTAAGAACAGGCTTCCATCCGTCGATCTTCATGCT 428
| | | | |
Db 244 TATGATACAGTGGTCACTTACACTTAAGAACAGGCTTCCATCCGTCGATCTTCATGCT 303
| | | | |
QY 429 GTTGGTGTATCTTACTGAGAAAGCTTCTGAGGAGCTGAATATGATGATGACAGACATCAA 488
| | | | |
Db 304 GTTGGTGTATCTTACTGAGAAAGCTTCTGAGGAGCTGAATATGATGATGACAGACATCAA 363
| | | | |
QY 489 AGGAGAAAGAGATATATAAGTCTTCCCTGGTGGAGCCATCATATATGCTGGCAGGTC 548
| | | | |
Db 364 AGGAGAAAGAGATATATAAGTCTTCCCTGGTGGAGCCATCATATATGCTGGCAGGTC 423
| | | | |
QY 549 CTGAAGAGAAATGGTCCATAGGCTTGAAGCCACAGCTGCTTACCTACGATCATATCTTCT 608
| | | | |
Db 424 CTGAAGAGAAATGGTCCATAGGCTTGAAGCCACAGCTGCTTACCTACGATCATATCTTCT 483
| | | | |
QY 609 CATGTGGACCTGGTAAAGAACTTGAATTCAGGCTTATGAGCCCTACTAGATATAGA 668
| | | | |
Db 484 CATGTGGACCTGGTAAAGAACTTGAATTCAGGCTTATGAGCCCTACTAGATATAGA 543
| | | | |
QY 669 GAAGGAGATCTGGCCACAGAAAGACACAGACCTTGCACAAATTTATCTACTTTTGGCT 728
| | | | |
Db 544 GAAGGAGATCTGGCCACAGAAAGACACAGACCTTGCACAAATTTATCTACTTTTGGCT 603
| | | | |
QY 729 GTATTGTATGAAGGAAAGTTGGACATGAGAAACAAAGACTCTTGTATGAGATAGG 788
| | | | |
Db 604 GTATTGTATGAAGGAAAGTTGGACATGAGAAACAAAGACTCTTGTATGAGATAGG 663
| | | | |
QY 789 GATGCTGATCTGCTCGGCTTGGCTTAAATGACACAGCTCATATGATATGAAACAG 848
| | | | |
Db 664 GATGCTGATCTGCTCGGCTTGGCTTAAATGACACAGCTCATATGATATGAAACAG 723
| | | | |
QY 849 TCTGCGCAGGCTGATTTGGATGACACAGAAATCAGTATATGGATGATGATG 908
| | | | |
Db 724 TCTGCGCAGGCTGATTTGGATGACACAGAAATCAGTATATGGATGATGATG 783
| | | | |
QY 909 GGCACCACTCTGAGAGTCAATATCTCGAAGGTCAACATTTCTTGGAGAAC 968
| | | | |
Db 784 GGCACCACTCTGAGAGTCAATATCTCGAAGGTCAACATTTCTTGGAGAAC 843
| | | | |
QY 969 CATGCCAGGCTCTTGGAAATCTGCGCAATTAATCTTCTTACTGCTCAACACTCTTG 1028
| | | | |
Db 844 CATGCCAGGCTCTTGGAAATCTGCGCAATTAATCTTCTTACTGCTCAACACTCTTG 903
| | | | |
QY 1029 ATGGACCTTGGACAGTTCTACTGTTTGTCTATCTCTCCACCAACATATGSCATG 1088
| | | | |
Db 904 ATGGACCTTGGACAGTTCTACTGTTTGTCTATCTCTCCACCAACATATGSCATG 963
| | | | |
QY 1089 GAAGCTTATGTCAAAGTACAGCTGTCCAGAGAACCCCACTACGAAATGAATATAT 1148
| | | | |
Db 964 GAAGCTTATGTCAAAGTACAGCTGTCCAGAGAACCCCACTACGAAATGAATATAT 1023
| | | | |
QY 1149 GAAGAGGAGGAGACTATATGATGATCTTACTGATTTCTGAATGAGATGGTCAAGTTT 1208
| | | | |
Db 1024 GAAGAGGAGGAGACTATATGATGATCTTACTGATTTCTGAATGAGATGGTCAAGTTT 1083
| | | | |
QY 1209 GATGATGACACTCTCTCTTTCCTTATCCAAATTTGGCTCAGTGGCCAAAGAGATCTTAA 1266
| | | | |
Db 1084 GATGATGACACTCTCTCTTTCCTTATCCAAATTTGGCTCAGTGGCCAAAGAGATCTTAA 1143
| | | | |
QY 1269 ACTTGGGTACTTACATTTGCTGCTGAAGAGAGAGACTGGGATGATGCTTACTGCTC 1328
| | | | |
Db 1144 ACTTGGGTACTTACATTTGCTGCTGAAGAGAGAGACTGGGATGATGCTTACTGCTC 1203
| | | | |
QY 1329 GCCCCCGATGACAGAGATTATTAAGTCAATATTTGAACAATGGCCCTCAGCGGATTTGGT 1388
| | | | |
Db 1204 GCCCCCGATGACAGAGATTATTAAGTCAATATTTGAACAATGGCCCTCAGCGGATTTGGT 1263
| | | | |
QY 1389 AGGAAGTCAAAAAAGTCCGATTTATGSCATACAGAGATGAACCTTTAAAGCTGTGAA 1448
| | | | |
Db 1264 AGGAAGTCAAAAAAGTCCGATTTATGSCATACAGAGATGAACCTTTAAAGCTGTGAA 1323
| | | | |
QY 1449 GCTATTCAGATGATCAGGAATCTTGGACCTTACTTATGGGAGATTTGAGACACA 1508
| | | | |

Db 1324 GCTATTCAGCATGAATTCAGGAATCTTGGACCTTACTTATTTGGGAGATTTGAGACACA 1383
| | | | |
QY 1509 CTGTGATTAATTTAATCAATCAAGCAGACCATTAATCATCTACCCACAGGAATC 1568
| | | | |
Db 1384 CTGTGATTAATTTAATCAATCAAGCAGACCATTAATCATCTACCCACAGGAATC 1443
| | | | |
QY 1569 ACTGATGTCCGCTCTTGTATTCAGAGAGATTTACCAAAAGGTGTAACATTTGAAGAT 1628
| | | | |
Db 1444 ACTGATGTCCGCTCTTGTATTCAGAGAGATTTACCAAAAGGTGTAACATTTGAAGAT 1503
| | | | |
QY 1629 TTTCCAAATTCGACAGGAGAAATTTCAATTAATTAATGACAGTACTGTAGAAAGTGG 1688
| | | | |
Db 1504 TTTCCAAATTCGACAGGAGAAATTTCAATTAATTAATGACAGTACTGTAGAAAGTGG 1563
| | | | |
QY 1689 CCACTAATCAGATCCCTGGGTGCTGACCGGCTTATCTACTTTCTGTAATATGAG 1748
| | | | |
Db 1564 CCACTAATCAGATCCCTGGGTGCTGACCGGCTTATCTACTTTCTGTAATATGAG 1623
| | | | |
QY 1749 AGAGATCTAGCTTCAGACATATTTGGCCCTCTCCATCTGCTACAAAGAAATCTGATAGT 1808
| | | | |
Db 1624 AGAGATCTAGCTTCAGACATATTTGGCCCTCTCCATCTGCTACAAAGAAATCTGATAGT 1683
| | | | |
QY 1809 CAAGAGAGAAACAGATATATCTCAGACAGAGAGATGTCATCTGTTTCTGTATTTGAT 1868
| | | | |
Db 1684 CAAGAGAGAAACAGATATATCTCAGACAGAGAGATGTCATCTGTTTCTGTATTTGAT 1743
| | | | |
QY 1869 GAGAACCCAGACTGGTATCTACAGAGAAATATACAGGCTTTCGCCAATCCAGCTGGA 1928
| | | | |
Db 1744 GAGAACCCAGACTGGTATCTACAGAGAAATATACAGGCTTTCGCCAATCCAGCTGGA 1803
| | | | |
QY 1929 GTGCACTTGGAGATTCAGAGTTCCAGGCTTCCAAATCATCATGACACATCATGCTAT 1988
| | | | |
Db 1804 GTGCACTTGGAGATTCAGAGTTCCAGGCTTCCAAATCATCATGACACATCATGCTAT 1863
| | | | |
QY 1989 GTTTTGAATGTTGGAGTTGCTGAGTTTGTTCATAGAGTGGCATCTGCTATCTA 2048
| | | | |
Db 1864 GTTTTGAATGTTGGAGTTGCTGAGTTTGTTCATAGAGTGGCATCTGCTATCTA 1923
| | | | |
QY 2049 AGCATTTGAGCAGACAGTCACTTCTTCTGCTTCTTCTGATATATACCTTCAACAC 2108
| | | | |
Db 1924 AGCATTTGAGCAGACAGTCACTTCTTCTGCTTCTTCTGATATATACCTTCAACAC 1963
| | | | |
QY 2109 AAAATGGTCTATGAAGACACATCACCCTATTTCCATTTCTCAGAGAGAACTGCTTATG 2168
| | | | |
Db 1984 AAAATGGTCTATGAAGACACATCACCCTATTTCCATTTCTCAGAGAGAACTGCTTATG 2043
| | | | |
QY 2169 TCGATGGAGAAACCCAGGCTATGATGATCTGGGTCGACAACTCAGACTTGGAGACAGA 2228
| | | | |
Db 2044 TCGATGGAGAAACCCAGGCTATGATGATCTGGGTCGACAACTCAGACTTGGAGACAGA 2103
| | | | |
QY 2229 GGCATGACCCGCTTACGAGGTTTCTAGTTGTGACAAAGAACTGGTGAATTTATGAG 2288
| | | | |
Db 2104 GGCATGACCCGCTTACGAGGTTTCTAGTTGTGACAAAGAACTGGTGAATTTATGAG 2163
| | | | |
QY 2289 GACGATTAAGAGATATTTAGCATCTTCTGAGTAAAGAAATGECATTTGAACCAAGA 2348
| | | | |
Db 2164 GACGATTAAGAGATATTTAGCATCTTCTGAGTAAAGAAATGECATTTGAACCAAGA 2223
| | | | |
QY 2349 AGCTTCTCCAG----- 2361
| | | | |
Db 2224 AGCTTCTCCAGAGATCTCTTGTGGGATTAACCACTATGTACTAGATACCAAGAA 2283
| | | | |
QY 2361 ----- 2361
| | | | |
Db 2284 GAGTGAATCCCAAGAGAGTCAACCAAGAAAAACAGCTTTTAAGAAAAAGATACATTT 2343
| | | | |
QY 2361 ----- 2361
| | | | |
Db 2344 TTGTCCCTGAACGCTTGTGAAGCAATCATGATAGCAGCAATTAATGAGGACAAAT 2403
| | | | |
QY 2361 ----- 2361
| | | | |

Db 2404 AAGCCGAATAGAGTCACTGGGCAAGCAAGGTAGAGTGAAGGCTGTCTCTCAA 2463
Qy 2381 AACCCACAGTCTTGAAGCCATCAAGGGAAATACCTGCTACTCTTCACTACAT 2420
Db 2464 AACCCACAGTCTTGAAGCCATCAAGGGAAATACCTGCTACTCTTCACTACAT 2523
Qy 2421 CAAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2480
Db 2524 CAAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2583
Qy 2481 ATTATGATGAGGATGAAATATGAGCCCGCCAGCTTTCAGAAAGAAACAGACATAT 2540
Db 2584 ATTATGATGAGGATGAAATATGAGCCCGCCAGCTTTCAGAAAGAAACAGACATAT 2643
Qy 2541 TTTATGCTGAGTGAAGAGCTCTGGGATTAATGAGATGATGATGATGATGATGAT 2600
Db 2644 TTTATGCTGAGTGAAGAGCTCTGGGATTAATGAGATGATGATGATGATGATGAT 2703
Qy 2601 AGAAGCAGGCTCAGAGTGGAGTGGCTCCCTCAGTCAAGAAAGTGTTCAGGAATTT 2660
Db 2704 AGAAGCAGGCTCAGAGTGGAGTGGCTCCCTCAGTCAAGAAAGTGTTCAGGAATTT 2763
Qy 2661 ACTGATGCTCCTTACTACAGCCCTTATACCTGGAGAACTAAATGAACATTTGGGACTC 2720
Db 2764 ACTGATGCTCCTTACTACAGCCCTTATACCTGGAGAACTAAATGAACATTTGGGACTC 2823
Qy 2721 CTGGGGCCATATATAGAGCAGAAATGAGATTAATATCATGCTTCACTTCAAGAAATCAG 2780
Db 2824 CTGGGGCCATATATAGAGCAGAAATGAGATTAATATCATGCTTCACTTCAAGAAATCAG 2883
Qy 2781 GCCCTGCTCCCTATCTCTCTATCTAGCCCTTATTTCTTATGAGAGATGAGAGGCA 2840
Db 2884 GCCCTGCTCCCTATCTCTCTATCTAGCCCTTATTTCTTATGAGAGATGAGAGGCA 2943
Qy 2841 GGAGCAGAACTAGAAAAAACTTGTCAAGCCTTAATGAAGCCAACTACTTTTGGAA 2900
Db 2944 GGAGCAGAACTAGAAAAAACTTGTCAAGCCTTAATGAAGCCAACTACTTTTGGAA 3003
Qy 2901 GTGCAACATCATATGGCAGCCCTTAAAGTGAATGATGATGATGATGATGATGATGAT 2960
Db 3004 GTGCAACATCATATGGCAGCCCTTAAAGTGAATGATGATGATGATGATGATGATGAT 3063
Qy 2961 TCTGATGTTGACCTGGAAAAAGATGAGCACTGAGGCTGATGAGACCCCTTGTGCTGC 3020
Db 3064 TCTGATGTTGACCTGGAAAAAGATGAGCACTGAGGCTGATGAGACCCCTTGTGCTGC 3123
Qy 3021 CACACTAACACACTGACCTGCTCATGGAGACAAAGTACAGTACAGAAATTTGCTCTG 3080
Db 3124 CACACTAACACACTGACCTGCTCATGGAGACAAAGTACAGTACAGAAATTTGCTCTG 3183
Qy 3081 TTTTTCACCATCTTGTGATGAGACCAAAAGCTGTACTTCACTGAAATATGGAAGAAAC 3140
Db 3184 TTTTTCACCATCTTGTGATGAGACCAAAAGCTGTACTTCACTGAAATATGGAAGAAAC 3243
Qy 3141 TGCAGGGCTCCCTGCAATATCCAGATGGAAGTCCCACTTTAAAGAGAAATATGCTTC 3200
Db 3244 TGCAGGGCTCCCTGCAATATCCAGATGGAAGTCCCACTTTAAAGAGAAATATGCTTC 3303
Qy 3201 CATGCAATCAATGCTACATATGATGATGATGATGATGATGATGATGATGATGATGAT 3260
Db 3304 CATGCAATCAATGCTACATATGATGATGATGATGATGATGATGATGATGATGATGAT 3363
Qy 3261 AGGATTCATGCTCTCTCAGCATGAGGCAATGAAACATCCATCTTATTCATTTTC 3320
Db 3364 AGGATTCATGCTCTCTCAGCATGAGGCAATGAAACATCCATCTTATTCATTTTC 3423
Qy 3321 AGTGAACATGCTCTCTCAGCATGAGGCAATGAAACATCCATCTTATTCATTTTC 3380
Db 3424 AGTGAACATGCTCTCTCAGCATGAGGCAATGAAACATCCATCTTATTCATTTTC 3483
Qy 3381 TATCCAGGTGTTTTGAGACAGTGAAGTGTACATCCAAAGCTGGAATTTGGCGGTG 3440
Db 3484 TATCCAGGTGTTTTGAGACAGTGAAGTGTACATCCAAAGCTGGAATTTGGCGGTG 3543

Qy 3441 GAATGCTTATTTGGGACATCTACATGCTGGGATGACACACTTTTCTGTGTACAGC 3500
Db 3544 GAATGCTTATTTGGGACATCTACATGCTGGGATGACACACTTTTCTGTGTACAGC 3603
Qy 3501 AATAAGTGTGAGTCCCTGGGAAATGCTTGTGACACATATGATTTTCAGATTACA 3560
Db 3604 AATAAGTGTGAGTCCCTGGGAAATGCTTGTGACACATATGATTTTCAGATTACA 3663
Qy 3561 GCTTCAGCAATATGACAGTGGGCCCCAAAGCTGGCCAGACTTATTTCCGGATCA 3620
Db 3664 GCTTCAGCAATATGACAGTGGGCCCCAAAGCTGGCCAGACTTATTTCCGGATCA 3723
Qy 3621 ATCAATGCTGTGAGCAGCAGAGAGCCCTTTTCTTGTGATCAAGTGTGATGCTGACCA 3680
Db 3724 ATCAATGCTGTGAGCAGCAGAGAGCCCTTTTCTTGTGATCAAGTGTGATGCTGACCA 3783
Qy 3681 ATGATTATTCAGGGATCAAGACCCAGGGTGGCGGTGAGAGTTCTCCAGCCTTACATC 3740
Db 3784 ATGATTATTCAGGGATCAAGACCCAGGGTGGCGGTGAGAGTTCTCCAGCCTTACATC 3843
Qy 3741 TCTCAGTTATCATCATATATGATGATGATGATGATGATGATGATGATGATGATGAT 3800
Db 3844 TCTCAGTTATCATCATATATGATGATGATGATGATGATGATGATGATGATGATGAT 3903
Qy 3801 TCCACTGCAACCTTATGCTCTTCTTGGCAATGATGATGATGATGATGATGATGATGAT 3860
Db 3904 TCCACTGCAACCTTATGCTCTTCTTGGCAATGATGATGATGATGATGATGATGATGAT 3963
Qy 3861 ATTTTAAACCTCCCAATATGCTGATACATCGCTTGAACCCCAACATATATGCAAT 3920
Db 3964 ATTTTAAACCTCCCAATATGCTGATACATCGCTTGAACCCCAACATATATGCAAT 4023
Qy 3921 CGCAGACCTCTTCGATGAGTGTGATGCTGTGATTTAAATAGTTGACAGATGCCATTG 3980
Db 4024 CGCAGACCTCTTCGATGAGTGTGATGCTGTGATTTAAATAGTTGACAGATGCCATTG 4083
Qy 3981 GGAATGAGAGTAAAGCAATATCAGATCAGACATGATGCTTATCTTACAT 4040
Db 4084 GGAATGAGAGTAAAGCAATATCAGATCAGACATGATGCTTATCTTACAT 4143
Qy 4041 ATGTTGCACTGCTCTCCCTCAAAAGCTGACCTCACTCCAGGGAGAGTAAATGCC 4100
Db 4144 ATGTTGCACTGCTCTCCCTCAAAAGCTGACCTCACTCCAGGGAGAGTAAATGCC 4203
Qy 4101 TGGAGACCTCAGGTAATATCCAAAGAGTGGCTGCAAGTGGACTTCCAGAGCAATG 4160
Db 4204 TGGAGACCTCAGGTAATATCCAAAGAGTGGCTGCAAGTGGACTTCCAGAGCAATG 4263
Qy 4161 AAAGTACAGAGTACTACTCAGGAGTAAATCTGCTTACAGCATGATGTGAAG 4220
Db 4264 AAAGTACAGAGTACTACTCAGGAGTAAATCTGCTTACAGCATGATGTGAAG 4323
Qy 4221 GAGTTCCTATCTCCAGAGTCAAGATGCCATGATGATGATGATGATGATGATGATGAT 4280
Db 4324 GAGTTCCTATCTCCAGAGTCAAGATGCCATGATGATGATGATGATGATGATGATGATGAT 4383
Qy 4281 AAAGTAAAGGTTTTTCAAGGAAATCAAGATCTTCAACAGCTGTGGTAACTCTTACAG 4340
Db 4384 AAAGTAAAGGTTTTTCAAGGAAATCAAGATCTTCAACAGCTGTGGTAACTCTTACAG 4443
Qy 4341 CCACCGTTACTGCTGCTACCTTGAATTCACCCCAAGAGTGTGGTCCACAGATTGCC 4400
Db 4444 CCACCGTTACTGCTGCTACCTTGAATTCACCCCAAGAGTGTGGTCCACAGATTGCC 4503
Qy 4401 CTGAGGATGAGGTTCTGGCTGCGAGGACAGAGACCTCTTAC 4442
Db 4504 CTGAGGATGAGGTTCTGGCTGCGAGGACAGAGACCTCTTAC 4545

RESULT 11
N81545
ID N81545 standard; DNA; 4616 BP.

AC N81545: (first entry)
DE 04-DEC-1990 Human Factor VIII-C analog having exon 4 and 14 deleted.
AW Human Factor VIII-C analog; exon deletion; coagulation disorders;
KW hemophilia; ss.
PN EP-265778-A.
PD 04-MAY-1988.
PE 14-OCT-1987: 115043.
PR 15-OCT-1986: US-919153.
PA (RORE) Rorer Int Overseas.
PI Saver N. D'ohan W.;
DR WPI: 88-120930/18.
P1 Human Factor VIII-C analogue free of other proteins - produced from
PT recombinant DNA and used in the treatment of coagulation disorders.
PS Claim 4: Page 13: 42pp. English.
CC The genetically engineered analogue can provide a dependable and
readily available therapeutic agent to be used in the treatment of
hemophilia and coagulation disorders in humans or animals.
See also N81543-45.
Sequence 4616 BP: 1333 A: 1038 C: 1012 G: 1232 T:
Query Match 80.2%; Score 3877.2; DB 1; Length 4616;
Best Local Similarity 90.4%; Pred. No. 0;
Matches 4366; Conservative 1; Mismatches 9; Indels 454; Gaps 3;
QY 72 ATGCAATAGAGCTTCACCTGCTTCTTCTTCTGCTTTGCGATTCGCTTAGTGC 131
DB 1 ATGCAATAGAGCTTCACCTGCTTCTTCTTCTGCTTTGCGATTCGCTTAGTGC 60
QY 132 ACCAGAGATACCTACTGCTGGTGCAGTGAAGTGCATGCGACTATGCAAGTATC 191
DB 61 ACCAGAGATACCTACTGCTGGTGCAGTGAAGTGCATGCGACTATGCAAGTATC 120
QY 192 GGTGAGTGCCTGTGGACGCAAGATTTCTCTAGAGTGCCTTTTCATTCAC 251
DB 121 GGTGAGTGCCTGTGGACGCAAGATTTCTCTAGAGTGCCTTTTCATTCAC 180
QY 252 ACCTGAGTCTGTACAAAGAGCTGTGTGTGAATTCAGATCCCTTTTCAACATC 311
DB 181 ACCTGAGTCTGTACAAAGAGCTGTGTGTGAATTCAGATCCCTTTTCAACATC 240
QY 312 GCTAAGCAAGGCGACCTGGATGGGTCTGTAGTCTTCCATCCAGCTGAGGTTAT 371
DB 241 GCTAAGCAAGGCGACCTGGATGGGTCTGTAGTCTTCCATCCAGCTGAGGTTAT 300
QY 372 GATACAGTGTCTATTACACTTAGAAGATGGCTTCCATCTCTGAGTCTTCA 431
DB 301 GATACAGTGTCTATTACACTTAGAAGATGGCTTCCATCTCTGAGTCTTCA 360
QY 432 GGTGTATCTCTACTGAAAGCTTCTGAGGAGTGAATATGATGATGACCAATCAAG 491
DB 361 GGTGTATCTCTACTGAAAGCTTCTGAGGAGTGAATATGATGATGACCAATCAAG 388
QY 492 GAGAAAGAGATATAAAGTCTTCCCTGCTGGAAGCAATATGTCTGAGGCTCTG 551
DB 388 ----- 388
QY 552 AAAGAGATGTCATGCGCTCTGACCACTGTGCTTACCTACTCATATCTTTCTCAT 611
DB 388 ----- 388
QY 612 GTGACCTGTGAAGACTTGAATTCAGGCTCATTTGAGGCTTACTAGATGTAGAGA 671
DB 388 ----- 388
QY 672 GGGAGTCTGGCCAGAGAAAGACAGACCTTGACAAATTATATCTTTTGGCTGA 731
DB 388 GGGAGTCTGGCCAGAGAAAGACAGACCTTGACAAATTATATCTTTTGGCTGA 447
QY 732 TTGTATGAAGGAAAGTGTGGCACTGAGAAAGAAAGAACTCTTGTATGAGATAGGAT 791
DB 448 TTGTATGAAGGAAAGTGTGGCACTGAGAAAGAAAGAACTCTTGTATGAGATAGGAT 506

QY 792 GCTGATCTGCTGGGCGCTGGCCCTAAATGACACAGATCAATGTTATGTAACAGGTCT 851
DB 507 GCTGATCTGCTGGGCGCTGGCCCTAAATGACACAGATCAATGTTATGTAACAGGTCT 566
QY 852 CTGCCAGGTGTGATTTGATGTCACAGAGAAATCATGCTATTTGGCATGTGATGGGC 911
DB 567 CTGCCAGGTGTGATTTGATGTCACAGAGAAATCATGCTATTTGGCATGTGATGGGC 626
QY 912 ACCACTCTGAACTGACATCAATATTCCTGGAAGGTCACACATTTCTGTGAGAACAT 971
DB 627 ACCACTCTGAACTGACATCAATATTCCTGGAAGGTCACACATTTCTGTGAGAACAT 686
QY 972 CGCCAGGCTCTTGGAAATCTCGCAATTAATTTCTTACTGCTCAACACTCTTGAT 1031
DB 687 CGCCAGGCTCTTGGAAATCTCGCAATTAATTTCTTACTGCTCAACACTCTTGAT 746
QY 1032 GACCTGGACAGTTTCTACTGTTTGTGATATCTCTCCACCAACATGATGGCATGAA 1091
DB 747 GACCTGGACAGTTTCTACTGTTTGTGATATCTCTCCACCAACATGATGGCATGAA 806
QY 1092 GCTTATGTCAAGTAGACAGCTGTCCAGAGAAACCCCACTACGAATGAATATATGA 1151
DB 807 GCTTATGTCAAGTAGACAGCTGTCCAGAGAAACCCCACTACGAATGAATATATGA 866
QY 1152 GAAGCGAAGACTATGATGATGATCTTACTGATTTCTGAATGATGATGATGATGAT 1211
DB 867 GAAGCGAAGACTATGATGATGATCTTACTGATTTCTGAATGATGATGATGATGATGAT 926
QY 1212 GATGCAACTCTCTCTTATCCAAATTCCTGATGCTCCAGAGATCTTAAACT 1271
DB 927 GATGCAACTCTCTCTTATCCAAATTCCTGATGCTCCAGAGATCTTAAACT 986
QY 1272 TGGTATCATTAATGCTGCTGAAGAGAGAGTGGATGCTCCCTTACTCTCGGC 1331
DB 987 TGGTATCATTAATGCTGCTGAAGAGAGAGTGGATGCTCCCTTACTCTCGGC 1046
QY 1332 CCCGATGACAGATTTATTAAGTCAATTTTGAACAATGGCCCTCAGCGATTTGATG 1391
DB 1047 CCCGATGACAGATTTATTAAGTCAATTTTGAACAATGGCCCTCAGCGATTTGATG 1106
QY 1392 AAGTACAAAAGTCCGATTTATGTCATACAGATGAACTTTAAGCTCTGAAGCT 1451
DB 1107 AAGTACAAAAGTCCGATTTATGTCATACAGATGAACTTTAAGCTCTGAAGCT 1166
QY 1452 ATTCAGCATGAATCAGGAATCTTGGACCTTACTTTATGGGAAAGTTGAGACACTG 1511
DB 1167 ATTCAGCATGAATCAGGAATCTTGGACCTTACTTTATGGGAAAGTTGAGACACTG 1226
QY 1512 TTGATATATTTAATAATCAAGCAAGCAAGCAATTAATCAATCTACCTCAGCAATCA 1571
DB 1227 TTGATATATTTAATAATCAAGCAAGCAAGCAATTAATCAATCTACCTCAGCAATCA 1286
QY 1572 GATGTCGCTCTTGTATTCAGAGAGATTAACAAAAGGTGTAACATTTGAAGATTT 1631
DB 1287 GATGTCGCTCTTGTATTCAGAGAGATTAACAAAAGGTGTAACATTTGAAGATTT 1346
QY 1632 CCAATTTGCGAGAGAAATATTCATATATATGAGACAGTGTGAGAGATGGGCA 1691
DB 1347 CCAATTTGCGAGAGAAATATTCATATATATGAGACAGTGTGAGAGATGGGCA 1406
QY 1692 ACTAATAGATCTCGGTGCTGACCGCTATTAATCTAGTTTCTGTAATATGAGAGA 1751
DB 1407 ACTAATAGATCTCGGTGCTGACCGCTATTAATCTAGTTTCTGTAATATGAGAGA 1466
QY 1752 GATCAGCTCAGAGACTATGGGCGCTCTCTCATGCTCAAGAAATGTAAGTATGATCA 1811
DB 1467 GATCAGCTCAGAGACTATGGGCGCTCTCTCATGCTCAAGAAATGTAAGTATGATCA 1566
QY 1812 AGAGAAACAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1871
DB 1527 AGAGAAACAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1586

QY	1872	AAGCAACCTGCTGACCTCACAAGAAATATCAACGCTTCTCCCAATCCAGCTGAGATG	19313
QY	1872	AAGCAACCTGCTGACCTCACAAGAAATATCAACGCTTCTCCCAATCCAGCTGAGATG	19313
Db	1587	AACCGAACCTGTACTCCTCACAGAAATATCAACGCTTCTCCCAATCCAGCTGAGATG	16468
QY	1932	CAGCTTGAGATCCAGAGATCCAAAGCCTCCAAATCATGACAGCATCAATGGCTATGTT	19911
Db	1647	CAGCTTGAGATCCAGAGATCCAAAGCCTCCAAATCATGACAGCATCAATGGCTATGTT	17068
QY	1992	TTTGATAGTTTGCAGTTGTGACGTTTGTTCATGAGGTGGCATCTACTGTACTTCTAAGC	20511
Db	1707	TTTGATAGTTTGCAGTTGTGACGTTTGTTCATGAGGTGGCATCTACTGTACTTCTAAGC	17668
QY	2052	ATTGAGACAGCACTGACTCTCTTCTCTGTCTTCTCTGTGATATCTTCAACACAAA	21111
Db	1767	ATTGAGACAGCACTGACTCTCTTCTCTGTCTTCTCTGTGATATCTTCAACACAAA	18288
QY	2112	ATGCTATGAAAGACACACTCACCCATTCCTCCATCTCAGAGAAACGTCTTCATGCG	21711
Db	1827	ATGCTATGAAAGACACACTCACCCATTCCTCCATCTCAGAGAAACGTCTTCATGCG	18868
QY	2172	ATGGAACCCAGGCTCATGGAATCTGCGGTGCGACAACTCAGACTTCTGGAACAGAGC	22313
Db	1887	ATGGAACCCAGGCTCATGGAATCTGCGGTGCGACAACTCAGACTTCTGGAACAGAGC	19468
QY	2232	ATGACCCGCTTACTGAAAGGTTTCTAGTTGTGACAAACACACTGTGATATTATACAGAC	22911
Db	1947	ATGACCCGCTTACTGAAAGGTTTCTAGTTGTGACAAACACACTGTGATATTATACAGAC	20068
QY	2292	AGTATGAAAGATATTTAGATACACTGCTGAGTAAACAAATGSCATTGAACCAAGAC	23511
Db	2007	AGTATGAAAGATATTTAGATACACTGCTGAGTAAACAAATGSCATTGAACCAAGAC	20668
QY	2352	TTCTCCGAG-----TTCTCCGAG-----	23611
Db	2067	TTCTCCGAGTTCTACTATTGGAATCCCTGCTTGGGATTAACCATGATGCTACTAGATA	21288
QY	2361	-----TTCTCCGAGTTCTACTATTGGAATCCCTGCTTGGGATTAACCATGATGCTACTAGATA	23611
Db	2127	CCAAAGAGAGTGAATCCCAAGAGAAAGTACACAGAAAAACAGCTTTTAAGAAAAAG	21868
QY	2361	-----TTCTCCGAGTTCTACTATTGGAATCCCTGCTTGGGATTAACCATGATGCTACTAGATA	23611
QY	2187	GATACCATTTTGTCCTCGAAACGCTTGTGAAGCAATCATGCAATAGCAATTAATAG	22468
QY	2361	-----TTCTCCGAGTTCTACTATTGGAATCCCTGCTTGGGATTAACCATGATGCTACTAGATA	23611
QY	2247	GGACAAATTAAGCCCGAAATTAAGAAAGTCACTCGGCGCAACCAAGTAGACTGAAGCTG	23068
QY	2361	-----TTCTCCGAGTTCTACTATTGGAATCCCTGCTTGGGATTAACCATGATGCTACTAGATA	23611
QY	2307	TGCTCTCAAAACCCACAGCTTGTGAACGCCATCAACGGGAAATTAATCTGTACTACTT	23668
QY	2412	CAGTCAGATCAAGAGAAATTTGACTATGATGATACATATCACTGTAATTAAGAAAGAA	24268
Db	2367	CAGTCAGATCAAGAGAAATTTGACTATGATGATACATATCACTGTAATTAAGAAAGAA	24268
QY	2472	GATTTTGACATTTATGATGAGAGATGAATAACAGAGCCCCGAGCTTCAAAAAGAAACA	25311
Db	2427	GATTTTGACATTTATGATGAGAGATGAATAACAGAGCCCCGAGCTTCAAAAAGAAACA	24868
QY	2532	CGACACATATTTATTTGCTGCACTGAGAGAGGCTCTGGGATTTATGGATAGTAGTCCCA	25913
Db	2487	CGACACATATTTATTTGCTGCACTGAGAGAGGCTCTGGGATTTATGGATAGTAGTCCCA	25468
QY	2592	CATGCTTAAGAAACAGGGCTCAGAGGCGAGAGTCCCTCATGTTCAAGAAATGTGTTTC	26511
Db	2547	CATGCTTAAGAAACAGGGCTCAGAGGCGAGAGTCCCTCATGTTCAAGAAATGTGTTTC	26068
QY	2652	CAGGATTTACTGATGCTCTCTTACTCAGCCCTTATACCGTGGAGAACTTAATACAT	27111
Db	2607	CAGGATTTACTGATGCTCTCTTACTCAGCCCTTATACCGTGGAGAACTTAATACAT	26668
QY	2712	TTGGGACTCTGGGCGCATATATTAAGACGAAGTTGAAGATTAATCATCTGTAATCTTC	27711

Db	2667	TTGGGAACTCTGGGGCCCATATATAAGCAGGAAGTTGAAGATAATATCATGTGTAACCTTTC	2726
Qy	2772	AGAAATCAGGCGCTCTCGTCCCTATTCCTTATCTTAGCCTTATTTCTTATGAGAGAGAT	2833
Db	2727	AGAAATCAGGCGCTCTCGTCCCTATTCCTTATTCCTTAGCCTTATTTCTTATGAGAGAGAT	2786
Qy	2832	CAGAGCGAAGGAGAGAACTTAGAAAAAATTGTGCAGGCTTATGTAAACCAAACCTTAC	2891
Db	2787	CAGAGCGAAGGAGAGAACTTAGAAAAAATTGTCAAGCTTATGTAAACCAAACCTTAC	2846
Qy	2882	TTTGGAAAGTGCACATCATATGACACCCACTAAAGATGAGTTGACTGCAAGCCTGG	2955
Db	2847	TTTGGAAAGTGCACATCATATGAGCACCCCACTAAAGATGAGTTGACTGCAAGCCTGG	2906
Qy	2952	GCTTATTTCTGTGATGTGACCTGGAAAAAGATGTGCATCAGGCGCTGATGTGACCCCTT	3011
Db	2907	GCTTATTTCTGTGATGTGACCTGGAAAAAGATGTGCATCAGGCGCTGATGTGACCCCTT	2966
Qy	3012	CTGCTCTGCCACTTAACACACTGAACCCCTGCTCATGAGGAGACAAGTGACACTACAGAA	3077
Db	2967	CTGCTCTGCCACTTAACACACTGGAACCCCTGCTCATGAGGAGACAAGTGACACTACAGAA	3028
Qy	3072	TTTGTCTGTTTTCACCAATCTTTGATGAGACCAAAAGCTGGTACTTCACTGAAAATATG	3131
Db	3027	TTTGTCTGTTTTCACCAATCTTTGATGAGACCAAAAGCTGGTACTTCACTGAAAATATG	3086
Qy	3122	GAAGAAACCTCAGAGGCTCCCTCGAATATCCAAATGGAAGATCCCACTTTTAAAGAGAT	3199
Db	3087	GAAGAAACCTCAGAGGCTCCCTCGAATATCCAAATGGAAGATCCCACTTTTAAAGAGAT	3146
Qy	3192	TATCGCTTCATGCAATCAATGGCTACATATGATGATACACTACCTGGCTTACTAATAGCT	3255
Db	3147	TATCGCTTCATGCAATCAATGGCTACATATGATGATACACTACCTGGCTTACTAATAGCT	3206
Qy	3252	CAGGATCAAGAGATTCGATGATCTGTCTCAGCATGGGAGCAATGAAAAACATCATCTT	3311
Db	3207	CAGGATCAAGAGATTCGATGATCTGTCTCAGCATGGGAGCAATGAAAAACATCATCTT	3266
Qy	3312	ATTCAATTCACATGGAACATGTGTACATGTCGAAAAAAGAGAGATATAAATGGCACCTG	3377
Db	3267	ATTCAATTCACATGGAACATGTGTACATGTCGAAAAAAGAGAGATATAAATGGCACCTG	3328
Qy	3372	TACAATCTCTATCCAGAGTGTTTTGTGAGACAGTGGAAATGTACCATCCAAACCTGGAAT	3433
Db	3327	TACAATCTCTATCCAGAGTGTTTTGTGAGACAGTGGAAATGTACCATCCAAACCTGGAAT	3388
Qy	3432	TGGCGGGTGAATGCCCTTATTTGGGAGCATCTACATGCTGGATGAGCACACTTTTCTG	3491
Db	3387	TGGCGGGTGAATGCCCTTATTTGGGAGCATCTACATGCTGGATGAGCACACTTTTCTG	3446
Qy	3492	GTTTACACCAATATAGTGTACAGACTCCCCCTGGGAAATGGCTTTGTGACACTTATGAGATTT	3555
Db	3447	GTTTACACCAATATAGTGTACAGACTCCCCCTGGGAAATGGCTTTGTGACACTTATGAGATTT	3506
Qy	3552	CAGATTACAGCTTAGAGACAATATGAGACATGAGGCCCCCAAAGTGGCGAGACTTCATTAT	3611
Db	3507	CAGATTACAGCTTAGAGACAATATGAGACATGAGGCCCCCAAAGTGGCGAGACTTCATTAT	3566
Qy	3612	TCCGATCAATCAATGCGCTGGAGACCAAGAGACCCTTTCTTGTGATCAAGGTGATCTG	3677
Db	3567	TCCGATCAATCAATGCGCTGGAGACCAAGAGACCCTTTCTTGTGATCAAGGTGATCTG	3628
Qy	3672	TTGGACCAATGATTTATTCACGGCATTAAGACCAGAGGTGCCGTACGAAGTTCTCCACG	3733
Db	3627	TTGGACCAATGATTTATTCACGGCATTAAGACCAGAGGTGCCGTACGAAGTTCTCCACG	3688
Qy	3732	CTCTACATCTCTGAGTTTATATCATCATATAGCTTGATGGGGAAGAAGGCGAGACTTAT	3791
Db	3687	CTCTACATCTCTGAGTTTATATCATCATATAGCTTGATGGGGAAGAAGGCGAGACTTAT	3746
Qy	3792	CGAGAAATTCACATGGAACCTTAATGCTTTCTTGGCAATGTGGATTCATCTGGAGATA	3851

Db 3747 CGAGGAATTCACCTGGACCTTAATGTCCTTTGGCAATGTCATCTGGGATA 3806
 QY 3852 AACACAAATTTTTTAACCTCCATTAATGTCGTGATACATCGTTTGCACCCAACTCAT 3911
 Db 3807 AACACAAATTTTTTAACCTCCATTAATGTCGTGATACATCGTTTGCACCCAACTCAT 3866
 QY 3912 TATAGCATTCGACGACCTCTTCGAGAGAGTGATGGGCGTGATTAATAATAGTTGAGC 3971
 Db 3867 TATAGCATTCGACGACCTCTTCGAGAGAGTGATGGGCGTGATTAATAATAGTTGAGC 3926
 QY 3972 ATGCATTTGGGAATGGAGAGTAAGAATATACAGATCAGAGATTAATGCTTCACTTAC 4031
 Db 3927 ATGCATTTGGGAATGGAGAGTAAGAATATACAGATCAGAGATTAATGCTTCACTTAC 3986
 QY 4032 TTTACCAATATGTTGGCACCCTGCTCTTCAAAAGCTGACTTCACTTCAAGGAGG 4091
 Db 3987 TTTACCAATATGTTGGCACCCTGCTCTTCAAAAGCTGACTTCACTTCAAGGAGG 4046
 QY 4092 AGTAATGCTGGAGACCTCAGTGTAATATCAAAAGAGTGGTCAAGTGGACTTCAG 4151
 Db 4047 AGTAATGCTGGAGACCTCAGTGTAATATCAAAAGAGTGGTCAAGTGGACTTCAG 4106
 QY 4152 AAGACATGAAGTACACAGAGTAATCTACAGAGTAATAATCTGCTTACAGCATG 4211
 Db 4107 AAGACATGAAGTACACAGAGTAATCTACAGAGTAATAATCTGCTTACAGCATG 4166
 QY 4212 TATGTAGAGAGTTCCTCATCTCAGAGTCAAGATGCCATCAGTGGACTCTCTTTT 4271
 Db 4167 TATGTAGAGAGTTCCTCATCTCAGAGTCAAGATGCCATCAGTGGACTCTCTTTT 4226
 QY 4272 CAGAATGGCAAAATGAAGTTTTTTCAGGAAATCAAGACTCTTCAACCTGGTGGTAA 4331
 Db 4227 CAGAATGGCAAAATGAAGTTTTTTCAGGAAATCAAGACTCTTCAACCTGGTGGTAA 4286
 QY 4332 TCTCTAGACCCACCGTTACTGACTCGCTTCAATTCACCCCAAGTGGTGGTAA 4391
 Db 4287 TCTCTAGACCCACCGTTACTGACTCGCTTCAATTCACCCCAAGTGGTGGTAA 4346
 QY 4392 CAATATGCTGAGAGTGTCTGGGCTGGAGGCAAGAGACTCTTCAAGGTTGG 4451
 Db 4347 CAATATGCTGAGAGTGTCTGGGCTGGAGGCAAGAGACTCTTCAAGGTTGG 4406
 QY 4452 CCACTGAGACCTGCGACGCTGACCTCTCCCTCAGCTCCAGGGCAGTGGTCC 4511
 Db 4407 CCACTGAGACCTGCGACGCTGACCTCTCCCTCAGCTCCAGGGCAGTGGTCC 4466
 QY 4512 CCTGCTGCTCTTCACTCTTGTGCTAAATCTTGAAGACACTGCTTGAAGCTCTCTG 4571
 Db 4467 CCTGCTGCTCTTCACTCTTGTGCTAAATCTTGAAGACACTGCTTGAAGCTCTCTG 4526
 QY 4572 AATTACTATCATCAGCTGCTGATTTCTTGGTGGGGGCGCAGAGGTCATCCAAATT 4631
 Db 4527 AATTACTATCATCAGCTGCTGATTTCTTGGTGGGGGCGCAGAGGTCATCCAAATT 4586
 QY 4632 AACTTACTTACTTACTTCTTCTGACAGCTG 4661
 Db 4587 AACTTACTTACTTACTTCTTCTGACAGCTG 4616
 RESULT 12
 ID T69811
 AC T69811 standard: DNA; 5035 BP.
 DT 10-AUG-1997 (first entry)
 DE Factor VIII-dB695-HCII DNA.
 KW Factor VIII-dB695-HCII; heparin cofactor II; blood coagulation;
 KW blood clotting; procoagulant; anticoagulant; antithrombotic;
 KW haemophilia; gene therapy; ss.
 OS Homo sapiens.
 FH Location/Qualifiers
 FT key 35..5020
 FT /tag= a
 FT /product= Factor VIII-dB695-HCII

FT misc_rna 2225..2314
 FT /tag= b
 FT /product= heparin cofactor II region (aa51-81)
 PN M09718315-A1.
 PD 22-MAY-1997.
 PF 13-NOV-1996; E04977.
 PR 13-NOV-1995; US-558107.
 PA (IMMO) IMMUNO AG.
 PI Vooiberg JJ;
 DR MPI: 97-289291/26.
 DR P-PSDB: M18670.
 PT Hybrid Factor VIII with modified activity, comprises region from
 PT donor anticoagulant or antithrombotic protein - useful for treatment
 PT of coagulation disorders
 PS Claim 16; Page 52-60; 96pp; English.
 CC A DNA molecule (T69811) codes for Factor VIII-dB695-HCII (M18670),
 CC a hybrid protein in which amino acids 712-736 of Factor-dB695
 CC (Factor VIII del868-1562) B-domain are replaced by amino acids
 CC 51-80 from the acidic region (and potential thrombin-binding site)
 CC of human heparin cofactor II (HCII). It was obtd. by PCR
 CC amplification (see also T69812-13) of the HCII acidic region from
 CC total liver cDNA, fusion to sequences encoding Factor VIII
 CC aa706-711 and aa737-743, and incorporation of the construct into
 CC plasmid pCIB-dB695. The hybrid protein, which can be expressed
 CC using gene therapy techniques, has increased procoagulant activity
 CC owing to the HCII acidic region, and can be used to treat blood
 CC coagulation disorders such as haemophilia A.
 SQ Sequence 5035 BP; 1484 A; 1127 C; 1110 G; 1314 T;
 Query Match 76.5%; Score 3696.4; DB 1; Length 5035;
 Best Local Similarity 86.7%; Pred. No. 0;
 Matches 4362; Conservative 0; Mismatches 56; Indels 612; Gaps 2;
 QY 43 CTCACATTAACATTTGTAGCAATAAGTCAATAGAGCTTCACCTGCTTCTTC 102
 Db 6 CTCACATTAACATTTGTAGCAATAAGTCAATAGAGCTTCACCTGCTTCTTC 65
 QY 103 TGTGCTTTTGGCATTCCTTTAGTGCACCAAGAAATACCTGAGTGGAGGA 162
 Db 66 TGTGCTTTTGGCATTCCTTTAGTGCACCAAGAAATACCTGAGTGGAGGA 125
 QY 163 TGTGAGGAGACATATAGCAAGAGATCTGGTGGTGGTGGTGGTGGTGGTGGTGG 222
 Db 126 TGTGAGGAGACATATAGCAAGAGATCTGGTGGTGGTGGTGGTGGTGGTGGTGG 185
 QY 223 CTGAGTGGCAAAATTTTTCATTCACACCTCAGTCTGTACAAAGAGCTGTGG 282
 Db 186 CTGAGTGGCAAAATTTTTCATTCACACCTCAGTCTGTGTACAAAGAGCTGTGG 245
 QY 283 TAGAATTCACGGATCACCTTTTCAACATGCTAAGCAAGGCCACCTGGATGGTCTGC 342
 Db 246 TAGAATTCACGGATCACCTTTTCAACATGCTAAGCAAGGCCACCTGGATGGTCTGC 305
 QY 343 TAGGCTCTACATCCAGGCTGAGGTTATGATACAGTGTCTAATACCTAAGACATGG 402
 Db 306 TAGGCTCTACATCCAGGCTGAGGTTATGATACAGTGTCTAATACCTAAGACATGG 365
 QY 403 CTTCCTATCTGATGCTTCACTGCTGTGTATCTTCTACTGAAAGCTTCTGAGGAG 462
 Db 366 CTTCCTATCTGATGCTTCACTGCTGTGTATCTTCTACTGAAAGCTTCTGAGGAG 425
 QY 463 CTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 522
 Db 426 CTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 485
 QY 523 GAAGCATATCATATGCTGTGGAGTCTTCAAGAGATGTCATGAGGCTCTGACCCAC 582
 Db 486 GAAGCATATCATATGCTGTGGAGTCTTCAAGAGATGTCATGAGGCTCTGACCCAC 545
 QY 583 TGTGCTTACTACTATATATTTCTCATGTGAGCTGTAAAGACTGAATTCAGGCC 642
 Db 546 TGTGCTTACTACTATATATTTCTCATGTGAGCTGTAAAGACTGAATTCAGGCC 605

0Y	643	TCCTTGGAGCCCTCTAGTATGTATGAGAGAGGAGTGTGGCCACAGAAAGACACAGACT	702
Db	606	TCATTGGAGCCCTACCTAGTATGTATGAGAGAGGAGTGTGGCCACAGAAAGACACAGACT	665
0Y	703	TGCACAAATTTATCTACTTTTGTCTGTATTTTATGAAAGGAAAAAGTTGGCACTCAGAA	762
Db	666	TGCACAAATTTATCTACTTTTGTCTGTATTTTATGAAAGGAAAAAGTTGGCACTCAGAA	725
0Y	753	CAAAAGACTCCTTGATGCAGGATAGGGATGCTGCATCTGCTGGGCTGTGGCCTAAATGC	822
Db	726	CAAAAGACTCCTTGATGCAGGATAGGGATGCTGCATCTGCTGGGCTGTGGCCTAAATGC	785
0Y	823	ACACAGTCAAATGTTATGTAAACAGGCTCTGCGACGTGTGTTGGATGCCACAGAAAT	882
Db	786	ACACAGTCAAATGTTATGTAAACAGGCTCTGCGACGTGTGTTGGATGCCACAGAAAT	845
0Y	883	CAGTCTATTGGCAGTGTATGGATGGGCGCACCTCCTGAAATGCACTCAATATTCCTCG	942
Db	846	CAGTCTATTGGCAGTGTATGGATGGGCGCACCTCCTGAAATGCACTCAATATTCCTCG	905
0Y	943	AAGGTACACATTTCTTGTGAGAACCATCGCCAGCGTCTTGGAAATCTGGCCATTA	1002
Db	906	AAGGTACACATTTCTTGTGAGAACCATCGCCAGCGTCTTGGAAATCTGGCCATTA	965
0Y	1003	CTTTCCTACAGCTCAAAACACTGTGATGGACCTTGGAGCTTGGAGTTCTACTGTTTGTCTA	1062
Db	966	CTTTCCTACAGCTCAAAACACTGTGATGGACCTTGGAGCTTGGAGTTCTACTGTTTGTCTA	1025
0Y	1063	TCTCTCCACCAACATGATGGCATGGAACTTATGTCAAATAGACAGCTGTCCAGAG	1122
Db	1026	TCTCTCCACCAACATGATGGCATGGAACTTATGTCAAATAGACAGCTGTCCAGAG	1085
0Y	1123	AACCCCAACTACGATATATAATATATGAGAAAGCGAAGACTATGATGATCTTACTG	1182
Db	1086	AACCCCAACTACGATATATAATATATGAGAAAGCGAAGACTATGATGATCTTACTG	1145
0Y	1183	ATTCTGAAATGGATGTGTGATGGTTGATGATACAACTCTCTTCCTTATCCAAATTC	1242
Db	1146	ATTCTGAAATGGATGTGTGATGGTTGATGATACAACTCTCTTCCTTATCCAAATTC	1205
0Y	1243	GCTCAGTTGCCAAGAGCATCTTAAACCTTGGGTACATTACATTGCTGTGTGAAGAGAG	1302
Db	1206	GCTCAGTTGCCAAGAGCATCTTAAACCTTGGGTACATTACATTGCTGTGTGAAGAGAG	1265
0Y	1303	ACTGGGACTATGCTCCCTTAGTCTGCTGCCCCGATACAGAAAGTTATTAAGTCAATTT	1362
Db	1266	ACTGGGACTATGCTCCCTTAGTCTGCTGCCCCGATACAGAAAGTTATTAAGTCAATTT	1325
0Y	1363	TGAACAATGGCCCTCACACGATGTGTGAGAGTACAAAAGAGCCGATTATGGCATCA	1422
Db	1326	TGAACAATGGCCCTCACACGATGTGTGAGAGTACAAAAGAGCCGATTATGGCATCA	1385
0Y	1423	CAGATGAACCTTTAAGACTGTGAAGCTATTGCGATGAATCAGGAATCTTGGGACCTT	1482
Db	1386	CAGATGAACCTTTAAGACTGTGAAGCTATTGCGATGAATCAGGAATCTTGGGACCTT	1445
0Y	1483	TACTTATGGGGAAGTTGGAGACACACTGTGATTATATTTAAGATCAAGCAGACAGC	1542
Db	1446	TACTTATGGGGAAGTTGGAGACACACTGTGATTATATTTAAGATCAAGCAGACAGC	1505
0Y	1543	CATATTAACATTTACCCCTCACAGAAATCACTATGCTCGCTTGTATCAAGAGATATAC	1602
Db	1506	CATATTAACATTTACCCCTCACAGAAATCACTATGCTCGCTTGTATCAAGAGATATAC	1565
0Y	1603	CAAAAAGGTATAAAACATTTGAAGATTTTCCATTTCTGCCAGAGAGAAATATTCAAAATTA	1662
Db	1566	CAAAAAGGTATAAAACATTTGAAGATTTTCCATTTCTGCCAGAGAGAAATATTCAAAATTA	1625
0Y	1663	AATGACAGTACTGTGTGAAGATGGGCACTAAATCAGATCCTGTGGGTCCGACCGCT	1722
Db	1626	AATGACAGTACTGTGTGAAGATGGGCACTAAATCAGATCCTGTGGGTCCGACCGCT	1685

QY	1723	ATTACTCAGTTTCGTGTAATAATGGAGAAGATCTACGCTTCAAGAGCTCATTTGGCCCTCTCC	1782
Dd	1686	ATTACTCAGTTTCGTGTAATAATGGAGAAGATCTACGCTTCAAGAGCTCATTTGGCCCTCTCC	1745
QY	1783	TCACTCGCTACAAGAAATCTGTAGATCAAAGAGAAACAGATTAATGTGCACGAAGAGA	1842
Dd	1746	TCACTCGCTACAAGAAATCTGTAGATCAAAGAGAAACAGATTAATGTGCACGAAGAGA	1805
QY	1843	ATGTCACTCGTTTTCTGTAATTTTGATGAGAACCGAAGCTGTACCTCACAGAGATATATAC	1902
Dd	1806	ATGTCACTCGTTTTCTGTAATTTTGATGAGAACCGAAGCTGTACCTCACAGAGATATATAC	1865
QY	1903	AACGCTTTCTCCCATTCCACATCCAGCTGAGGTGACGCTTBSAGATCCAGAGTTCCAAAGCTCCA	1962
Dd	1866	AACGCTTTCTCCCATTCCAAACGAGCTGAGGTGACGCTTBSAGATCCAGAGTTCCAAAGCTCCA	1925
QY	1963	ACATCATGACACAGCATCAATAGGCTATGTTTGGATGTTGGAGTTGCAGTTGGTTTGC	2022
Dd	1926	ACATCATGACACAGCATCAATAGGCTATGTTTGGATGTTGGAGTTGCAGTTGGTTTGC	1985
QY	2023	ATGAGGTGGCATCTGGTATCTTCTAAGCATTTGSAGACAGACAGACTGACTTCTCTGTCT	2082
Dd	1986	ATGAGGTGGCATCTGGTATCTTCTAAGCATTTGSAGACAGACAGACTGACTTCTCTGTCT	2045
QY	2083	TCTTCTCTGGAATATACCTTTAAAACAAAAATGCTGTATGAGACACACTCACCTATTACC	2142
Dd	2046	TCTTCTCTGGAATATACCTTTAAAACAAAAATGCTGTATGAMACACACTCACCTATTACC	2105
QY	2143	CATTCTCGSAGAAACGTCTCTATGTCGATGSGAAACCCAGGCTCATGSAATTCGGGGT	2202
Dd	2106	CATTCTCGSAGAAACGTCTCTATGTCGATGSGAAACCCAGGCTCATGSAATTCGGGGT	2165
QY	2203	GCCACAACTCAGACTTTTCGGAGACAGAGGCTGACGCCCTTACTGAAAGTTCTTAGTTGT-	2262
Dd	2166	GCCACAACTCAGACTTTTCGGAGACAGAGGCTGACGCCCTTACTGAAAGTTCTTAGTTGT	2225
QY	2262	-----GACAGAACACTGCTGATTTATTACGAGACAGTTATGAGATATT	2307
Dd	2226	TTCCAGAGGGGAGAGAGACGAGACTATCTGGACCTCGAGAAATATTCAGTGAGAGAC	2285
QY	2308	CAGCATTTCTGTAGTAAAAACAATGSCATTGAACCAAGAAGCTTCCGCCAG-----	2361
Dd	2286	ACGACTCATCTGACATCGTCGACAGTCTGTATGTAAACCAAGAAAGCTTCCCGAAATTCAA	2345
QY	2361	-----	2361
Dd	2346	GACACCCGACACTAGGCAAAAGCAATTTAATGCCACACCATTCCAGAAAATGCAATAG	2405
QY	2361	-----	2361
Dd	2406	AGAAAGACTGACCTTGGTTTGACACAGAAACACCTATGCTTAATAATACAAAATGTCTCT	2465
QY	2361	-----	2361
Dd	2466	CTAGTATTTGTTGATGCTCTCTTGCGACAGAGTCTTACTCCACAGTGGGCTATCTTATCTG	2525
QY	2361	-----	2361
Dd	2526	ATCTCCAAAGAACCAAAATATGAGACTTTTTTCTGTATGATCCATCACCCTGGAGCAATAGCA	2585
QY	2361	-----	2361
Dd	2586	GTAATAACAGCTGTCTGAAATGACACACTTCAGGCCACAGCTCCATCAGATGGGGACA	2645
QY	2361	-----	2361
Dd	2646	TGATATTACCCCTGAGTCAAGGCTCCCAATTAGATTAAATGAGAAGTGGGACAACTG	2705
QY	2361	-----	2361
Dd	2706	CAGATCCTCTTGCTTGGGATTAACCACTATGTACTAGATACCAAAAGAAAGTGGAAAT	2765
QY	2361	-----	2361

Db 2766 CCCAAGAGAGAGTCCACGAGAAAAACAGCTTTTAGAAAAAGGATACCATTTGTCCCTGA 2825
QY 2361 ----- 2361
Db 2826 ACGCTTGTGAAGCAATCATGCAATAGCAGCAATAAATGAGGGACAAATAAAGCCGGAAA 2885
QY 2361 -----MAOCCACCG 2370
Db 2886 TAGAACTCACCTGGGCAAAAGCAAGTAGAGACTGMAAGGCTGTGCTCTCAAAACCCACCG 2945
QY 2371 TCTTGAAGGCCATCAACGGGAAATTAATCTGTAATCTCTGATGATGATCAAGAGAAA 2430
Db 2946 TCTTGAAGGCCATCAAGGGAAATTAATCTGTAATCTCTGATGATGATCAAGAGAAA 3005
QY 2431 TTGACTATGATGATACCATATCATGTTGAATGAGAGAGAAATTTGACATTATGATG 2490
Db 3006 TTGACTATGATGATACCATATCATGTTGAATGAGAGAGAAATTTGACATTATGATG 3065
QY 2491 AGGATGAAATGAGAGCCCGAGCTTCAAAAGAAACAGACATATTTATTTGCTG 2550
Db 3066 AGGATGAAATGAGAGCCCGAGCTTCAAAAGAAACAGACATATTTATTTGCTG 3125
QY 2551 CAGTGGAGAGGCTCTGGGATTTATGGGATGATAGCTCCACATGTTCTAAGAAACAGGG 2610
Db 3126 CAGTGGAGAGGCTCTGGGATTTATGGGATGATAGCTCCACATGTTCTAAGAAACAGGG 3185
QY 2611 CTCAGAGTGGCGAGTCCCTCCTCAGTTCAAGAAAGTTTTCAGAGAAATTTACTGATGCT 2670
Db 3186 CTCAGAGTGGCGAGTCCCTCCTCAGTTCAAGAAAGTTTTCAGAGAAATTTACTGATGCT 3245
QY 2671 CCTTACTCAGCCCTTATACCTGGAGAACTAATGAACATTTGGGACCTCGGGGGCAT 2730
Db 3246 CCTTACTCAGCCCTTATACCTGGAGAACTAATGAACATTTGGGACCTCGGGGGCAT 3305
QY 2731 AATAGAAGCAGAGTGAAGATAATCATGTAATCTTCAAGAAATCAGGCCCTCGCT 2790
Db 3306 AATAGAAGCAGAGTGAAGATAATCATGTAATCTTCAAGAAATCAGGCCCTCGCT 3365
QY 2791 CCTATCTCTCTATCTGATGAGCTTATTTCTTATGAGAGAGATCAGAGGCAAGGCAAC 2850
Db 3366 CCTATCTCTCTATCTGATGAGCTTATTTCTTATGAGAGAGATCAGAGGCAAGGCAAC 3425
QY 2851 CTAGAAAAAACTTGTCAAGGCTTAATGAACCAAACTTACTTTGAAAGGCAACATC 2910
Db 3426 CTAGAAAAAACTTGTCAAGGCTTAATGAACCAAACTTACTTTGAAAGGCAACATC 3485
QY 2911 AATAGGCAACCACTAAGATGAGTTGACTGCAAAAGCCTGGGCTTATTTCTGATGTTG 2970
Db 3486 AATAGGCAACCACTAAGATGAGTTGACTGCAAAAGCCTGGGCTTATTTCTGATGTTG 3545
QY 2971 AACTGGAAAAAGATGATGCACTCAGGCTGATGAGACCCCTTGTGCTGCACTAACA 3030
Db 3546 AACTGGAAAAAGATGATGCACTCAGGCTGATGAGACCCCTTGTGCTGCACTAACA 3605
QY 3031 CACTGAACCCCTGCTCAGTGGAGAGACAAGTACAGTACAGGAATTTCTGTTTTCACCA 3090
Db 3606 CACTGAACCCCTGCTCAGTGGAGAGACAAGTACAGTACAGGAATTTCTGTTTTCACCA 3665
QY 3091 TCTTTGATGAGACCAAAAGCTGTGTTCACTGAAATATGAGAAAGAACTGCAAGGCTC 3150
Db 3666 TCTTTGATGAGACCAAAAGCTGTGTTCACTGAAATATGAGAAAGAACTGCAAGGCTC 3725
QY 3151 CCTGCAATATCCAGATGAGAGATCCCATTTTAAAGAGATTAATGCTTCATGCAATCA 3210
Db 3726 CCTGCAATATCCAGATGAGAGATCCCATTTTAAAGAGATTAATGCTTCATGCAATCA 3785
QY 3211 ATGGCTACATATGATGATCACTACTGCTTAGTAATGGCTCAGGATGAAGAGATTCAT 3270
Db 3786 ATGGCTACATATGATGATCACTACTGCTTAGTAATGGCTCAGGATGAAGAGATTCAT 3845
QY 3271 GGTATCTGCTCAGATGGGAGCAATGAAAAATCATCTTATTCATTTGAGTGAGATG 3330
|||||

Db 3846 GGTATCTGCTCAGATGGGAGCAATGAAAAATCATCTTATTCATTTCAAGTGACATG 3905
QY 3331 TGTTCACGTACGAAAAAAGAGAGATTAATAATGGCACTGTACAAATCTATCCAGGTG 3390
Db 3906 TGTTCACGTACGAAAAAAGAGAGATTAATAATGGCACTGTACAAATCTATCCAGGTG 3965
QY 3391 TTTTGAACAGAGTGGAAATGTTTACATCCAAAGCTGCAATTTGGCGGGTGGAAATGCTTA 3450
Db 3966 TTTTGAACAGAGTGGAAATGTTTACATCCAAAGCTGCAATTTGGCGGGTGGAAATGCTTA 4025
QY 3451 TTGGGACATCTCATATGCTGGATGAGACACATTTTCTGCTGTACAGCAATAGTCTC 3510
Db 4026 TTGGGACATCTCATATGCTGGATGAGACACATTTTCTGCTGTACAGCAATAGTCTC 4085
QY 3511 AGACTCCCTGGGAATGCTTCTGACACATTAAGATTTTTCAGATTACAGCTTCAGAGAC 3570
Db 4086 AGACTCCCTGGGAATGCTTCTGACACATTAAGATTTTTCAGATTACAGCTTCAGAGAC 4145
QY 3571 AATATGACAGTGGGCCCCCAAGCTGGCAGACTTCATTAATCCGATCAATAGTCTC 3630
Db 4146 AATATGACAGTGGGCCCCCAAGCTGGCAGACTTCATTAATCCGATCAATAGTCTC 4205
QY 3631 GGAGCACCAGAGAGCCCTTTCTTGATCAAGTGGATGCTGTCGACCAATGATTTTC 3690
Db 4206 GGAGCACCAGAGAGCCCTTTCTTGATCAAGTGGATGCTGTCGACCAATGATTTTC 4285
QY 3691 ACGGATCAAGACCAAGGCTCCGTCAGAAAGTTCTCCAGCTCTACATCTCTCAGTTTA 3750
Db 4286 ACGGATCAAGACCAAGGCTCCGTCAGAAAGTTCTCCAGCTCTACATCTCTCAGTTTA 4325
QY 3751 TCATCATGATATAGTCTTGATGAGAAAGAGTGGAGACTTATCAGAGAAATTCACCTGAA 3810
Db 4326 TCATCATGATATAGTCTTGATGAGAAAGAGTGGAGACTTATCAGAGAAATTCACCTGAA 4385
QY 3811 CCTTAATGCTCTTTTGGCAATGATGATTCATCTGGGATTAACACAATATTTTAAC 3870
Db 4386 CCTTAATGCTCTTTTGGCAATGATGATTCATCTGGGATTAACACAATATTTTAAC 4445
QY 3871 CTCGAATTTATGCTGATACATCCGTTTGCAACCACTAATTAAGATTCGACGACTC 3930
Db 4446 CTCGAATTTATGCTGATACATCCGTTTGCAACCACTAATTAAGATTCGACGACTC 4505
QY 3931 TTGCGATGAGATGATGAGGCTGTATTAATAGTTGACAGATCCCAATGGGAATGAGA 3990
Db 4506 TTGCGATGAGATGATGAGGCTGTATTAATAGTTGACAGATCCCAATGGGAATGAGA 4565
QY 3991 GTAAAGCAATATCAGATGACAGATTAATGCTCATCTTCAATTAATGATGTTGCCA 4050
Db 4566 GTAAAGCAATATCAGATGACAGATTAATGCTCATCTTCAATTAATGATGTTGCCA 4625
QY 4051 CTTGGTCTCTTCAAAAGCTGACCTCACTCCAAGGAGAGATTAATGCTGAGACTC 4110
Db 4626 CTTGGTCTCTTCAAAAGCTGACCTCACTCCAAGGAGAGATTAATGCTGAGACTC 4685
QY 4111 AGGTGAATTAATCCAAAGAGTGGTGAAGTGGACTTCAGAAAGACATATAAGTCCAG 4170
Db 4686 AGGTGAATTAATCCAAAGAGTGGTGAAGTGGACTTCAGAAAGACATATAAGTCCAG 4745
QY 4171 GAGTAATCTACAGAGGAGTAAATCTCTGCTACACAGCATGTAATGTAAGAGATCTCA 4230
Db 4746 GAGTAATCTACAGAGGAGTAAATCTCTGCTACACAGCATGTAATGTAAGAGATCTCA 4805
QY 4231 TCTCCAGAGTCAAGATGGCCATCAGTGAATCTCTTTTTCAGAAATGGCAAGTAAGG 4290
Db 4806 TCTCCAGAGTCAAGATGGCCATCAGTGAATCTCTTTTTCAGAAATGGCAAGTAAGG 4865
QY 4291 TTTTTCAGGGAATCAAGATCTCTTCACTGCTGTGGAATCTCTAGACCAAGCTTAC 4350
Db 4866 TTTTTCAGGGAATCAAGATCTCTTCACTGCTGTGGAATCTCTAGACCAAGCTTAC 4925
QY 4351 TGACTGCTACCTTGAATTCACCCCAAGTGGGGTGGACCAAGATTTGCCCTGAGATG 4410
Db 4926 TGACTGCTACCTTGAATTCACCCCAAGTGGGGTGGACCAAGATTTGCCCTGAGATG 4985
|||||

1569 ACTGATGCGCTGTTGTATTCAGAGGAGATTACCAAGGTGTAATTAACATTTGAAGAT 1628
1500 ACTGATGTCAGCGCTTTGACCCAGGAGACTTCTAAAGGTTGAAACATTTGAAAGAC 1559
1629 TTTCATTTCTGCAGAGAGAAATATTCATTAATTAATGACAGTGCATGAGAGATGGG 1688
1560 ATGCAATTCCTGCAGAGAGAGACTTTCATATTAATGACAGTGCATGAGAGATGGG 1619
1689 CCACATTAATCAATCTCGGTGCTGACCCGCTATTACTCTGTTGTTGTTAATGAG 1748
1620 CCACCAAGTCCATCTCGGTGCTGACCCGCTACTACTGAGAGTCCATTAATCTAGAG 1679
1749 AGAGATCTACTGAGAGTCCATTTGGCCCTCTGCTCATCTGTCACAAAGATCTAGAT 1808
1680 AAGATCTGCTTCGGAGTCTTGGACCTTCTGCTCTGCTGTCACAAAGATCTGAGAC 1739
1809 CAAAGAGAAACAGATTAATGTCAGACAGAGAGATGTCATCTGTTTCTGTTATTTGAT 1868
1740 CAAAGAGAAACAGATGATGATGACAGAAAGAGTCACTGTTTCTGTTATTCAT 1799
1869 GAGAACCCGAACTGCTGCTCCTCAGAGAAATATACAAAGCTTTTCCCAATCCAGCTGGA 1928
1800 GAAATCAAGAGCTGCTGCTCCTCAGAGAAATATACAGGCTTCCCTCCCAATCCGAGTGA 1859
1929 GTGCACTTGGAGATCCAGATCCAGATCCAGAGCTCCAGATCATGACAGATCATGGCTAT 1988
1860 TTACAGCCCGAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGAT 1919
1989 GTTTTGTATAGTTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2048
1920 GTTTTGTATAGTTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1979
2049 AGCATTTGAGAGACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2108
1980 AGTGTGTGAGACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2039
2109 AATATGCTCATGAGACAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAG 2168
2040 AATATGCTCATGAGACAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAG 2099
2169 TCGATGAGAAACCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAG 2228
2100 TCAATGAGAAACCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAG 2159
2229 GGATATCCGCTTACTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2288
2160 GGATATCCGCTTACTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2219
2289 GACATTTATGAGATATTTGAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAG 2348
2220 AACCTTATGAGATATTTGAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAG 2279
2349 AGCTTCTCCAGAACCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAG 2408
2280 -----GACATTAAGCCCTTCTACT 2297
2409 CTTCAGTACATCAAGAGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2468
2298 TTTCAGCCCGAGAGAACAAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2357
2469 GAGATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2528
2358 GAGATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2417
2529 ACACGACTATTTTATTTGCTGAGTGGAGAGGCTTGGAGATTTGAGATGATGATGATGATGATG 2588
2418 ACCGACACTATTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2477
2589 CCACATGTTCTAAGAAACAGGCTGCTGAGTGGAGAGTCCCTCAGTTCAGTTCAGAAAGTTGTT 2648
2478 CCGCGGCGCTTGAAGAAACAGGCTGCTGAGAGGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCT 2537
2649 TTCAGAGATTTACTGATGCTCTTACTCAGCCCTTATACGCTGGAGAACTAATGAA 2708
2538 TTCCGGGAATTTGCTGACGGCTCTTACAGCGCCGCTGCTACCGGGGAACTCAACAA 2597
2709 CATTTGGAGCTCTGCGGCGCATATATAGAGAGAGTGAAGATTAATTCATGATGATGATGATG 2768
2598 CACTTGGGCTCTTGGAGACCTTACATCAGAGCGGAAGTTGAAGACATCATGATGATGATGATG 2657
2769 TTACAGATCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2828
2658 TTCAAAAACAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2717
2829 GATCAGAGGCAAGAGAGCAAACTAGAAAAAATTTGCAACCTTAATGAAACCAAACT 2888
2718 GATCAGAGGCAAGAGAGCAAACTAGAAAAAATTTGCAACCTTAATGAAACCAAACT 2777
2889 TACTTTGGAAGTCAATCATATATGAGACAGGCTTGAAGATGATGATGATGATGATGATGATG 2948
2778 TACTTTGGAAGTCAATCATATATGAGACAGGCTTGAAGATGATGATGATGATGATGATGATG 2837
2949 TGGCTTATTTCTGATGTTGACCTGGAAGAAAGTGTGACACTCAGGCTGATTTGAGACC 3008
2838 TGGCTTATTTCTGATGTTGACCTGGAAGAAAGTGTGACACTCAGGCTGATTTGAGACC 2897
3009 CTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3068
2898 CTTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2957
3069 GAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3128
2958 GAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3017
3129 ATGGAAGAACTGCAAGGCTCCCTGCAATATCAGATGGAAGATCCACTTTTAAAG 3188
3018 GTGGAAGAACTGCAAGGCTCCCTGCAATATCAGATGGAAGATCCACTTTTAAAG 3077
3189 AATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3248
3078 AATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3137
3249 GCTCAGATCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3308
3138 GCTCAGATCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3197
3309 TCTATTTATTTAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3368
3198 TCGATTTATTTAGGAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3257
3369 CTTCAGATCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3428
3258 GTTCAGATCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3317
3429 AATTTGCGGAGTGAATGCTTATTTGCGAGATGATGATGATGATGATGATGATGATGATGATGATG 3488
3318 AATTTGCGGAGTGAATGCTTATTTGCGAGATGATGATGATGATGATGATGATGATGATGATGATG 3377
3489 CTG 3548
3378 CTG 3437
3549 TTTCAGATTTAAGCTTCAAGGCAATATGAGAGAGGCGCCCAAGAGCTGCCAGATTTAT 3608
3438 TTTCAGATTTAAGCTTCAAGGCAATATGAGAGAGGCGCCCAAGAGCTGCCAGATTTAT 3497
3609 TATTCGAGATCAATCAATGCTGAGAGCAAGAGAGGCTTTTCTTGTGATCAAGGTGAT 3668
3498 TATTCGAGATCAATCAATGCTGAGAGCAAGAGAGGCTTTTCTTGTGATCAAGGTGAT 3557
3669 CTGTTGGCAACATGATTTATTCAGGCAATCAAGAGAGGCTTCCCTGATCAAGGTGAT 3728
3558 CTGTTGGCAACATGATTTATTCAGGCAATCAAGAGAGGCTTCCCTGATCAAGGTGAT 3617
3729 AGCTTCAATCTCTCAATTTATCATCATGATATAGTCTTGTGATGGAAGATGAGAGCT 3788

QY	1029	ATGACCTTGGACAGTTTCTTACTGTTTGTGCAATCTCTTCCACCACATGATGGCATG	1088
Dp	984	ATGGACCTTAGGCGCAGTTCTCTGCTCTTCTTGCCACATGACGAGCCACCGACACAGCGGCATG	1043
QY	1089	GAAAGTTATGTCAAAAGTAGACAGCTGTCCAGAGAAACCCCAACTACGAAATGAAATAAT	1148
Dp	1044	GAGGTTTACGTGAAGGTGAGACAGTGTGCCCGAGSAGGCCAGCTGTGGCATGAAGAACAAC	1103
QY	1149	GAAGAACGGGAAGACTATGATGATCTTACTGATTTGAAAATGAGTGTGGCAGGTTT	1208
Dp	1104	GAGGAGGCGCGAGGACTACGACAGACAGCCTACCCGACAGGAGATGATGTGCTAGCTTC	1163
QY	1209	GATGATGACACCTCTCTCTTATCCAAATTCGTCAGTTGGCCAAAGAAAGATCCTAA	1268
Dp	1184	GACGACGACCAACCCCCAGCTTATCCAAATCCGAGCGGTGGCCAAAGAACCTTAA	1223
QY	1269	ACTTGGGTACATTATCTGCTGTGAAGAGAGAGAGCTGGAGACTATCTCTTACTCTCTC	1328
Dp	1224	ACCTGGGTGACTACATCTGCGCGCGGAGAGAGAGACTGGGACTAGGCCCTGTACTACTG	1283
QY	1329	GCCCCGATGACAAAGTTATTTAAAGTCATTTTGAACAATGGCCCTTAGCGGATTTGCT	1388
Dp	1284	GCCCCGAGCGACCGACTACAAAGGCAATACCTTAACAAAGGCCCCACAGGCATCTGGC	1343
QY	1389	AGGAAGTCAAAAAAGTCCCATTTATGGCATACACAGATGAACCTTTAAGACTGTGA	1448
Dp	1344	CGCAAGTACAAAGAAAGGTGGCGTTATGGCTTACCCGAGAGACTTTCAAGACCCGGAG	1403
QY	1449	GCTATTACAGCATGAATACGAAATCTTGGACCTTTCCTTATATGGGAAAGTTGGACACA	1508
Dp	1404	GCCATTCAGCAGGATCCGGCTCTCTGGCCCCCTCTGTACGGGAGGTGGGCGACACC	1463
QY	1509	CTGTGATATATTTAAGAAATCAAGCAGACAGCAATTAACATCTACCTCAGCGAATC	1568
Dp	1464	CTGTGATACATCTTCAGAAACAGCAGCCAGAGGCTTACAACTCTACCCCGACGGCATC	1523
QY	1569	ACTGATGTCCGTCCTTGTATTCAGAGAGATTACAAAAGTGTAAACATTTGAAGAT	1628
Dp	1524	ACCCAGCGGCCCCCCCTGTACAGCGCGCGCTCCCAAGGGCGTGAAGCACTGAAGAAC	1583
QY	1629	TTTCCAAATCTGCCAGGAGAAATTAATCAATTAATTAATGACAGTGACTGAATAATGGG	1688
Dp	1584	TTCCCCACTCTGCCCCGGCGAGATCTTAAATGAACAGTGGACGTGACCTGAGGAGCGC	1643
QY	1689	CCAACTAAATCAGATCCTCGGTGCTGTACCCGCTACCGCTATCTAGTTCTCTTAATATGAG	1748
Dp	1644	CCCAACAAAGACGACCCCGGTGCTGTACCCGCTACTACAGCAGCTGTGTGAACATGAG	1703
QY	1749	AGAAATCTAGCTTAAAGACATCATTTGGCCCTCTCTCATCTGCTACAAAGAAATCTAAT	1808
Dp	1704	CGCAACCTGGCTCTCGGACGTGATGGCCCCCTCTATCTGTGTACAAAGAGAGCGCTGAC	1763
QY	1809	CAAAAGAGAAACCGAATTAATGTACAGAAAGAGAAATGATCCTCTGTTCTTATATTTGAT	1868
Dp	1764	CAGGCGGCGCAACGATCATGAGGAGCAACAGCAAGTATCTCTGTTCAGCGTGTTCAC	1823
QY	1869	GAGAACCGAAGCTGTGTCCTCACAGAGATATACAAAGCTTTCTCCCAATCCAGCTGGA	1928
Dp	1824	GAGAACCGCACTGTGTATCTGACCCGAGAACATCACAGCTTCTGTGCGCAACCCCGCTGGC	1883
QY	1929	GTCGAGCTTGAGATCCAGATGTCCAGAGCTCCCAATCATGATGACAGCATCAATGGCTAT	1988
Dp	1884	GTCGAGCTGGAAGATCCCGAGTTCCAGGCTCCAGGCAACATATGATGACACGATCAACGGCTAC	1943
QY	1989	GTTTTGATAGTTGACAGTTGTGAGTTTGTTCGATGAGGTGGCATACTGTACATTTCTA	2048
Dp	1944	GTTTTCACACAGCTCGACGTGAGCGTGTGCTCATGAGGTGGCTACTGATGATCTGTG	2003
QY	2049	AGCAATGAGACAGACTGACTTCCTTTCTGTCTTTCTCTGTGATATACCTTCAACAC	2108
Dp	2004	AGCAATGAGCGCCGACGACGATCTCTGAGGTGTCTTCTCTCGGAGATACCTTCAAGAC	2063

QY	2109	AAAGGCTCTATGAAAGACACATCCCTATTCCATTCTCAGAGAAACTGTCTTATG	2168
Db	2064	AAGAGTGGTACGAGGACACCTCGACCTGTTCCTCTCCGGGAGACTGTGTTCTATG	2123
QY	2169	TGATGAGAAACCCAGAGTCTATGGATTCTGGGTGGCCAACTCAGACTTGTGGAAACGA	2228
Db	2124	TCTATGAGAAACCCCGGCTGTGGATTCTGGGTGGCCAACTCAGACTTGTGGAAACGC	2183
QY	2229	GGCATGACCGCCTTACTGAAGGTTTCTAGTTGTGACAAAGACACTGTGATATTACGAG	2288
Db	2184	GGCATGACTGCCCTGTGAAAGTCTCCAGCTGGACAAAGACCGCGGACTACTAGAG	2243
QY	2289	GACACTTATGAAAGTATTTCAGCACTCTGCTGATGAAAAAACAATGCCATTGACCAAGA	2348
Db	2244	GACACTCTCGAGGACATCTCCGCTACTGTGTGTCCAAAGAACAGCCATCGAGCCCGC	2303
QY	2349	AGCTTCTCCAGAA-----CCCA	2366
Db	2304	TCTTCTTCCAAAACCTCCCGCACCCAGCAGCGCTCAGAGCAAGTTCAGCGCACCCCC	2363
QY	2367	CCAGTCTTGAACCGCATCAACGGGAAATTAACTCGTACTCTTCAGTCAGATCAAGAG	2426
Db	2364	CCCGGTGTAAGGCGCACACGCGGAGATACCCGACACCCCTGCAAAAGGACCAAGAG	2423
QY	2427	GAATTAAGTCTTGTATGATTAACCATATCACTTTGAAATGAGAGGAAGATTTCACATTTAT	2486
Db	2424	GAGATCGACTCTCGACGACACCATCAGGTGAGATGAAAGAGACTTCGACACTTCAC	2483
QY	2487	GATGAGATGAAATTCAGAGCCCCCGAGTTTCAAAAGAAACAGACACTATTATT	2546
Db	2484	GACGAGAGACGAAACCGAGACCCCCCTCTTCCAAAAGAAACCCGCGCACTACTTCATC	2543
QY	2547	GCTGCAGTGAAGAGCTCTGGGATTATGGGATAGTAGCTCCCACTATTCTTAAGAAAC	2606
Db	2544	GCGCGCGTGAAGCGCCTGTGGGACTAGGAGATGACGAGCAGCCCAAGTCTGTGGCAAC	2603
QY	2607	AGGCTCAGAGTGGCACTGTCCCTCAGATTCAAGAAAGTTGTTTCCAGGAATTAAGTAT	2666
Db	2604	GCGCGCGAGACGGCACGTCGCCCACTTCAGAAAGTGTGTCCAGAGATTCAACGCAC	2663
QY	2667	GSCCTCTTATCAGCCTTATACGGTGAAGAACTAAATGAACATTGGGACTCCTGGGG	2726
Db	2664	GGCACTTACCCACGCCCTGTACCGGGGAGACTGAAGACACCTGTGGCTGTCTGGC	2723
QY	2727	CCATATATAGAGCAGAGATTGAAGATTAATATCATGTGTAACCTTCAGAAATCAGCCTCT	2786
Db	2724	CCCTACATCCGCGCGGAGGTGAGGAACAATATGTATCACTTCGCAACCAACCTCC	2783
QY	2787	GTCTCCATTCTCTTATCTTACGCTTATTTCTTATGAGGAAGATCAAGGCAAGAGCA	2846
Db	2784	GCGGCTACTCTCTTACTCTCCCTCCATGACGTACGAGAGGAACACGCGCAGGGCGCC	2843
QY	2847	GAACTGAAAAAACTGTGTCAGGCTAATGAAACCAAACTTACTTTTGGAAAGTGA	2906
Db	2844	GAGCCCGCAAGAACTTCGGAAGCCCAACGAGACTTAAGACTCTTGGAAAGGTGAG	2903
QY	2907	CATCATATGACACCCACTAAAGATGATGATGATGCAAGAGCTGTGGCTTATTTCTCAT	2966
Db	2904	CACACATAGGCGCCCAACAAAGAGAGATTCGATGCAAGCCTGGGCTACTTACGAGAC	2963
QY	2967	GTTGACTCGAAAAAAGATGTGCACTAAGGCTGATGTAACCCCTTGTGTGTCACACT	3026
Db	2964	GTTGACTCGGGAAGGACGTGCAACAGGCTGATGTAAGCCCTCTGTGTGTCACAC	3023
QY	3027	AACACACTGAACCTGCTCATGTGGAGACAAAGTGAAGTACAGGATTTTGTCTGTCTTC	3086
Db	3024	AACACCTGAACCCCGCCACGGGAGGAGTGTGTCAAGAAATTTGCCCTGTCTTC	3083
QY	3087	ACCATCTTGAATGAGCAAAAGCTGTGCTTCACTGAAATTAATGGAAGAACTGCAGG	3146
Db	3084	ACCATCTTTCAGCAGACTTAAGACTGTGTACTTCCAGGAACATGGAACGACATGCGGC	3143
QY	3147	GCTCCCTGCATATCCAGATGGAAGATCCCACTTTTAAAGAAATTAATGCTTCATCA	3206

Db	3144	GGCCCCGCAACATCCAGATGGAAGATCCCACTTCAAGGAGAATACCGCTTCCAGCGC	3203
Qy	3207	ATCAATGGCTACATAAATGGATACACTACCTAGGCTTAGTATGGCTCAGGATCAAAAGAT	3266
Db	3204	ATCAACGGCTACATATGAGAACACCTCCGCGCTGGTATGGCCAGGACCAAGGCGATC	3265
Qy	3267	CGATGGTATCTGCTCAGCATGGCGAGCATGAAGAACATCCATCTTATTCATTCAATGGA	3328
Db	3264	CGCTGGTACCTGCTGCTATGGCGAGCAGAGAAACATCCACAGATCCACTTCCAGCGGC	3323
Qy	3327	CATGTTCACTGTACGAAAAAAGAGAGATATAAATGGCACTGTACAAATCTTATGCA	3388
Db	3324	CACGTTTACCGCGCCAGAAAGAGAGATACAAAGATGGCCCTGTACAACTGTACACC	3383
Qy	3387	GGTGTTTTGAAGACATGGAATATTTACATCCAAAGCTGGAATTTGGGGGTGGAATGC	3448
Db	3384	GGCGGTTCGAGAGATGTGGAGATCTGCCAGCAAGCCGGATCTGGCGCTGGAGTGC	3443
Qy	3447	CTTATTTGGCGAGCATCTACATGCTGGAGTAGCACACTTTTCTGGTGTACAGCAATAG	3508
Db	3444	CTGATCGCGCAGACCTGTACAGCGCGGATAGCACCTGTGTTCTGGTGTACAGCAAG	3503
Qy	3507	TGTAGACTCCCGGGGAATGGCTCTCGAGACATTAAGATTTTCAATTTACACTTCA	3566
Db	3504	TGCCAGACCCCTGGGCAATGGCCAGCGGCACATCCGGACTTCCAGATCCACCCAGC	3563
Qy	3567	GGACAAATAGACAGTAGTGGGCCCAAGAGCTGGCAGACTTATTTATCCGATCAAT	3628
Db	3564	GGCCAGTAGCGGCAGTGGGCTCCCAAGCTGGCCGCTGACTACAGCGCGAGCATCAC	3623
Qy	3627	GCTTGGAGCACCAAGAGCCCTTTCTTGAGTCAAGTGAATCTGTTGGCACATGAT	3688
Db	3624	GCTGTGTGACAAAGAGCCCTTCTCTGGATCAAGTGAAGCTGTGCTGGCCCATGATC	3683
Qy	3687	ATTACGGCATCAAGACCAGGGGTGCCGTCAABAATTTCTCAGGCTCTACATCTCTAG	3746
Db	3684	ATCCAGGCAATCAAGACCAGCGGGGCCCGCAGAAATTCAGAGCTGTACATCCAGCAG	3743
Qy	3747	TTTATCATCATGTATTACTCTTGAATGGGAAGATGGCAGCACTTATCGAAGAAATCCACT	3806
Db	3744	TTTATCATCATGTATCTCTTGAAGGCGAAGATGGCAGAACTTACCGCGGCAACGCAAC	3803
Qy	3807	GGAACCTTAATGCTCTTTGGCAATGTGGATTCATCTGGATTAACACAAATATTTT	3866
Db	3804	GGCAACCTGAATGGTCTTCTTGGAAGGTGGAAGAGGGGATCAAGGACAAATCTTC	3863
Qy	3867	AACCTTCCATTAATGCTGCATACATCCGTTTGCAACCAATCATTAATAGCAATTCGACG	3926
Db	3864	AACCCCCCATCAATCGCCCTCAATCGCGCTGCACCCACCACTACAGCATCCGCGAGC	3923
Qy	3927	ACTTTGGCATGAGTGGATGGGCTGTGATTAATTAATGTGACATGGCATTTGGGAATG	3986
Db	3924	ACCTTGGCATGAGCTGATGGGCTGTGACATGAACAGTGTGAGATGGCCCTGGCGATG	3983
Qy	3987	GAGAGTAAGCAATATCAGATGCGACAGATTACTGCTTCACTTACATTAACCAATTTGTT	4046
Db	3984	GAGAGCAAGGCAATCAGCGACGCGCGAGATCAACGGCTCCAGGTACTTCCACAAATGTTC	4043
Qy	4047	GCCACCTGTGCTCTTCAAAAAGCTCGACTTCACCTCCAAGSAGAGATATGCTGGAGA	4106
Db	4044	GCCACCTGTGAGCCCGACAGAGCCCGCTGTCACTTCAGAGGCGCGAGCAACGCTCGCGC	4103
Qy	4107	CCTAGGTGATATATCCAAAAGATGGGTCGACAGTGGATTCACAAAGCAATGAANAATC	4166
Db	4104	CCCCAGGTGAACAACCCCAAGAGTGGCTCGAGGTGATTCACAAAAAACCATGAAGGTG	4163
Qy	4167	ACAGAGATACACTACTCAGGGAGTAAATCTCTGCTTACCAAGCATGTATGTGAAGAGTTC	4226
Db	4164	ACTGGCGTAGACACCAACCGGCGTCAAGAGCGTCTGACAGCATGTAGGTGAAGAGTTC	4223
Qy	4227	CTCATCTCCAGCAATCAAGATGGCAATCAATGAGCTCTCTTTTTCAGAAATGCAAGTA	4286

ID	Accession	Gene	Protein	Length	Score	DB	Length	Score	DB
Db	4224	CTGATCAGCAGACACCGAGAGGCGCACAGCACTGAGACCCCTGTCTTCCAAAACGGCAAGTG		4283					
Qy	4287	AAGGTTTTCAGGGAAATTCAGAACTCCTTCACACCTGTGGTGAATCTCTAGACCCACCG		4348					
Db	4284	AAGGTGTTCCAGGGCAACACAGAGAGCTTCACACCGGTGTGTAAGAGCTGGACCCGCC		4343					
Qy	4347	TTACTGACTGCTCACTTCGAAATTCACACCCCAAGTTGGGTGACACAGATTGCCCTGAGG		4406					
Db	4344	CTGCTGACCCGCTCACTGCGCGATTCACACCCCAAGTGGGTGACACAGATTGCCCTGCGC		4403					
Qy	4407	ATGAGGTTTCTGGGCTCCAGAGCAGACAGAGACTCTACTGAGGGTGGC		4452					
Db	4404	ATGAGAGTGTCTGGGCTCGAGAGCCCAAGAGACTGTACTGAAAGCGGCC		4449					
RESULT 15									
Q50185	050185	standard; cdNA; 9009 BP.							
AC	050185:								
DT	05-MAY-1994	(first entry)							
DE	Human factor VIII CDNA.								
KW	Blood clotting; coagulation; hybrid human/porcine factor VIII;								
KM	anti-haemophilic factor; factor 8; ss.								
OS	Homo sapiens.								
FH	Key	Location/Qualifiers							
FT	misc_feature	1..2277							
FT	/*tag= a	/note= "Domain structure: equivalent to the							
FT	Al-A2 domain"								
FT	5001..7053								
FT	/*tag= b	/note= "Domain structure: equivalent to the							
FT	A3-C1-C2 domain"								
FT	208..7206								
FT	cds								
FT	/*tag= c	/product= Factor_VIII							
FT	/note= "encodes R43257"								
PN	W09320093-A.								
PD	14-OCT-1993.								
PF	07-APR-1993.								
PR	07-APR-1992; US-864004.								
PA	(OYEM-) UNIV EMORY.								
PI	Loillar JS; Runge MS;								
DR	WPI: 93-336824/42.								
DR	P-PSDB: R43257.								
PT	Hybrid human and porcine factor VIII - having high coagulant								
PT	activity, used for treating patients with factor VIII deficiency								
PT	partic. haemophilia								
PS	Disclosure; Page 10-18; 94pp; English.								
CC	Hybrid human/porcine factor VIII molecules are claimed. These can be								
CC	obtained by (1) substituting a porcine subunit (i.e. heavy or light								
CC	chain) for the corresp. human subunit; (2) substituting a porcine								
CC	domain (i.e. A1, A2, A3, B, C1 and C2) for the corresp. human domain;								
CC	(3) substituting part of a porcine domain for the corresp. human								
CC	domain or (4) changing residues in the human sequence to the corresp.								
CC	porcine residues. The hybrid factor VIII molecules have greater								
CC	clotting activity than human factor VIII and are more resistant to								
CC	factor VIII inhibitors.								
SO	Sequence 9009 BP; 2853 A; 1907 C; 1844 G; 2405 T;								
Query Match									
		51.1%;	Score 2468.8;	DB 1;	Length 9009;				
	Best Local Similarity	99.7%;	Pred. No. 0;						
	Matches 2473; Conservative	0;	Mismatches	7;	Indels				

2465 GAAGAGATTTTACATTTATGATGAGTGAATTCAGAGCCCCGGACGTTTCAAA 2524
|||||
5226 GAAGAGATTTTACATTTATGATGAGTGAATTCAGAGCCCCGGACGTTTCAAA 5285
|||||
2525 GAAACACGACACTATTTTATGCTCAGTGGAGAGCTCTGGATTAATGGATGATAG 2584
|||||
5286 GAAACACGACACTATTTTATGCTCAGTGGAGAGCTCTGGATTAATGGATGATAG 5345
|||||
2585 CTTCCCAACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTCTCCCTCAGTCAAGAAAGT 2644
5346 CTTCCCAACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTCTCCCTCAGTCAAGAAAGT 5405
|||||
2645 TGTTCCTCAGGATTTACTGATGCTCTTACTCAGACCTTATACCGTGGAGACTAA 2704
|||||
5406 TGTTCCTCAGGATTTACTGATGCTCTTACTCAGACCTTATACCGTGGAGACTAA 5465
|||||
2705 TGAACATTTGGGACTCTGGGGCATATATAGAGAGAAATGAAGTAAATATCATGGT 2764
5466 TGAACATTTGGGACTCTGGGGCATATATAGAGAGAAATGAAGTAAATATCATGGT 5525
|||||
2765 AACTTCAGAAATCAGGCTCTGCTCCTTATTCCTTATTCAGCTTATTTCTTATGA 2824
|||||
5526 AACTTCAGAAATCAGGCTCTGCTCCTTATTCCTTATTCAGCTTATTTCTTATGA 5585
|||||
2825 GGAAGATCAGAGGCAAGGACAGACCTGAAAAACTTGTCAAGCTTAATGAACCA 2884
|||||
5586 GGAAGATCAGAGGCAAGGACAGACCTGAAAAACTTGTCAAGCTTAATGAACCA 5645
|||||
2885 AACTTACTTTGGAAAGTGCACATATAGCACCACCTAAAGATGATTTACTGCAA 2944
5646 AACTTACTTTGGAAAGTGCACATATAGCACCACCTAAAGATGATTTACTGCAA 5705
|||||
2945 AGCTGGGCTTATTTCTGATGTTTACCTGGAAAAAGATGTCACAGGCTGATGG 3004
5706 AGCTGGGCTTATTTCTGATGTTTACCTGGAAAAAGATGTCACAGGCTGATGG 3064
|||||
3005 ACCCTCTGCTGTCGACACTACACACTGAAACCTGCTCATGGAGAGCAAGTGCAGT 3064
5766 ACCCTCTGCTGTCGACACTACACACTGAAACCTGCTCATGGAGAGCAAGTGCAGT 5825
|||||
3065 ACAGGAATTTGCTCTGTTTTCACCATCTTTGATGAGAGCCAAAAGTGTACTTCACTGA 3124
5826 ACAGGAATTTGCTCTGTTTTCACCATCTTTGATGAGAGCCAAAAGTGTACTTCACTGA 5885
|||||
3125 AATATGGAAGAAATGCGAGGCTCCCTGCAATATCCAGATGGAAGTCCCACTTTAA 3184
5886 AATATGGAAGAAATGCGAGGCTCCCTGCAATATCCAGATGGAAGTCCCACTTTAA 5945
|||||
3185 AGAGATTTATCGCTTCATGCAATGAGCTACATATGATGATACACTACCTGGTTAGT 3244
5946 AGAGATTTATCGCTTCATGCAATGAGCTACATATGATGATACACTACCTGGTTAGT 6005
|||||
3245 AATGCTCAGATCAAAAGATGATGATGATGCTCAGAGATGGGCAATGAAACAT 3304
6006 AATGCTCAGATCAAAAGATGATGATGATGCTCAGAGATGGGCAATGAAACAT 6065
|||||
3305 CCATTCATTCATTTAGTGGACATGTTGCTACGTACGAAAAAAGAGAGATATAAAT 3364
6066 CCATTCATTCATTTAGTGGACATGTTGCTACGTACGAAAAAAGAGAGATATAAAT 6125
|||||
3365 GGCACGTACATCTCTATCCAGGCTTTTGGACAGTGGAAATTTCCATCCCAAGC 3424
6126 GGCACGTACATCTCTATCCAGGCTTTTGGACAGTGGAAATTTCCATCCCAAGC 6185
|||||
3425 TGAATTTGGCGGATGGAATGCTTATTTGGCGAGATCTACATGCTGGATGAGCACT 3484
6186 TGAATTTGGCGGATGGAATGCTTATTTGGCGAGATCTACATGCTGGATGAGCACT 6245
|||||
3485 TTTTCTGCTGACAGCAATAGTGTACAGCTCCCTGGGAAATGGCTTGGACACATTAAG 3544
6246 TTTTCTGCTGACAGCAATAGTGTACAGCTCCCTGGGAAATGGCTTGGACACATTAAG 6305
|||||
3545 AGATTTTACGATTACAGCTTACAGGACATATGACAGTGGGCCCAAGCTGGCCAGACT 3604

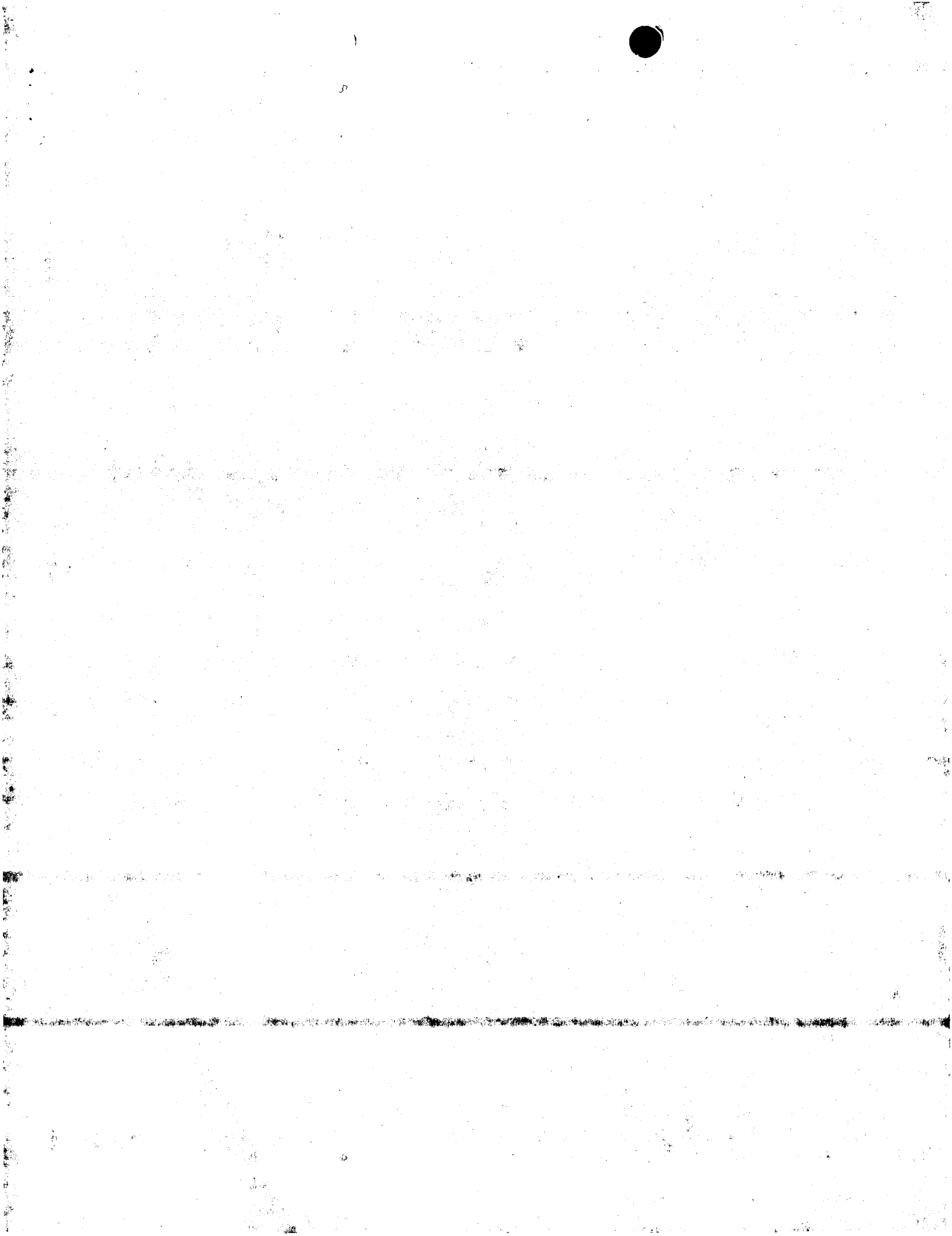
|||||
6306 AGATTTTACGATTACAGCTTACAGGACATATGACAGTGGGCCCAAGCTGGCCAGACT 6365
|||||
3605 TCATTAATTCGGATCAATCAATGCTGGAGCACCAAGAGCCCTTTCTTGGATCAAGT 3664
6366 TCATTAATTCGGATCAATCAATGCTGGAGCACCAAGAGCCCTTTCTTGGATCAAGT 6425
|||||
3665 GATCTGTTGGACCAATGATTTATCAGGGCATCAAGACCCAGGGTGGCCCTCAGAAAGT 3724
6426 GATCTGTTGGACCAATGATTTATCAGGGCATCAAGACCCAGGGTGGCCCTCAGAAAGT 6485
|||||
3725 CTTCAAGCTCTACATCTCAGATTATCATGATATAGTCTTGAAGGGAAGAAGGGA 3784
6486 CTTCAAGCTCTACATCTCAGATTATCATGATATAGTCTTGAAGGGAAGAAGTGGCA 6545
|||||
3785 GACTATTCAGGAAATTCACATGGAACCTTAATGGCTCTTGGCAATGATGATTCATC 3844
6546 GACTATTCAGGAAATTCACATGGAACCTTAATGGCTCTTGGCAATGATGATTCATC 6605
|||||
3845 TGGGTAACCAACATATTTTAACTTCCAATTTATGCTGATACATCCGTTGCACCC 3904
6606 TGGGTAACCAACATATTTTAACTTCCAATTTATGCTGATACATCCGTTGCACCC 6665
|||||
3905 AACTCATTTACGATTCGAGACCTTCGCAATGGAGTGGATGGGCTGATTTAAATAG 3964
6666 AACTCATTTACGATTCGAGACCTTCGCAATGGAGTGGATGGGCTGATTTAAATAG 6725
|||||
3965 TTGACAGATGCTCATTGGGATGAGAGTAAAGCAATATCAGATGACAGATGATGCTTC 4024
6726 TTGACAGATGCTCATTGGGATGAGAGTAAAGCAATATCAGATGACAGATGATGCTTC 6785
|||||
4025 ATCTCACTTTACCAATATGTTTGGCACTGGCTCTCTTCAAAAGTGCATTCACCTCCA 4084
6786 ATCTCACTTTACCAATATGTTTGGCACTGGCTCTCTTCAAAAGTGCATTCACCTCCA 6845
|||||
4085 AGGAGAGATTAATGCTGGAGACCTCAGGTGAATATCCAAAGAGTGGCTCAGTGA 4144
6846 AGGAGAGATTAATGCTGGAGACCTCAGGTGAATATCCAAAGAGTGGCTCAGTGA 6905
|||||
4145 CTTCCAGAAAGACAAATGAAGTACAGAGATTAATCAGAGGAGTAAATCTCTGCTTAC 4204
6906 CTTCCAGAAAGACAAATGAAGTACAGAGATTAATCAGAGGAGTAAATCTCTGCTTAC 6965
|||||
4205 CAGCATGATGTTGAAGGATTCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGCATCT 4264
6966 CAGCATGATGTTGAAGGATTCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGCATCT 7025
|||||
4265 CTTTTCAGATGCAAGTAAAGTTTTTCAGGAAATCAAGACTCCCTCAGACCTGT 4324
7026 CTTTTCAGATGCAAGTAAAGTTTTTCAGGAAATCAAGACTCCCTCAGACCTGT 7085
|||||
4325 GGTGAACCTCTCAGACCAACCTTACTGACTGCTACCTTCGAAATTCACCCCAAGATTG 4384
7086 GGTGAACCTCTCAGACCAACCTTACTGACTGCTACCTTCGAAATTCACCCCAAGATTG 7145
|||||
4385 GGTGACACAGATGCTCCTGAGATGAGGTTCTGGGCTGCGAGGACAGACCTTACTG 4444
7146 GGTGACACAGATGCTCCTGAGATGAGGTTCTGGGCTGCGAGGACAGACCTTACTG 7205
|||||
4445 AGGGTGGCCACTGAGACACCTGCGACCTGCGCCTCTCCCTCCTCAGCTCCAGGGGAG 4504
7206 AGGGTGGCCACTGAGACACCTGCGACCTGCGCCTCTCCCTCCTCAGCTCCAGGGGAG 7265
|||||
4505 TGTCCCTCCTGCTGCTGCTTCTACCTTGTGCTAAATCTTACAGACACTCCTTGAAG 4564
7266 TGTCCCTCCTGCTGCTGCTTCTACCTTGTGCTAAATCTTACAGACACTCCTTGAAG 7325
|||||
4565 CCTCCTGAATTAATCATCATGCTCTGCATTTCTTGGTGGGGGCGACAGAGGCTGAT 4624
7326 CCTCCTGAATTAATCATCATGCTCTGCATTTCTTGGTGGGGGCGACAGAGGCTGAT 7385
|||||
4625 CCAATTTAACTTAACCTTAATCTTCTGAGAGTGCCTCCAGATTACTCTCTCTCC 4684


```

Db 7386 CCAATTACTTAAGCTTACTTAATTTCTGCAGCTGCTCCAGATTACTTCCTTCC 7445
QY 4685 AATATTAAGTAAAGAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 4744
    |||||||
Db 7446 AATATTAAGTAAAGAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 7505
    |||||||
QY 4745 GTTAGGCTCTCAGAGTACACCTTCCCTGTGTAGAAAGAAAGAAAGAAAGAAAG 4804
    |||||||
Db 7506 GTTAGGCTCTCAGAGTACACCTTCCCTGTGTAGAAAGAAAGAAAGAAAGAAAG 7565
    |||||||
QY 4805 AAAAGATATTTATGATGTT 4824
    |||||||
Db 7566 AAAAGATATTTATGATGTT 7585
    |||||||

```

Search completed: August 18, 1999, 14:52:44
 Job time: 1392 sec



GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

nm nuc1eic - nucleic search, using sw model

Run on: August 18, 1999, 14:29:31 ; Search time 758.83 Seconds

(without alignments)
12560.497 Million cell updates/sec

Title: US-09-001-039a-46

Sequence: 1 CTCGAGCTAAAGATATTTTA.....ATTATGATGTCGGCCGCC 4832

Scoring table: IDENTITY_NUC

Searched: 2546578 seqs, 986266752 residues

Database :

1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: em_est20: *
21: em_est21: *
22: em_est22: *
23: em_est23: *
24: em_est24: *
25: em_est25: *
26: em_est26: *
27: em_est27: *
28: em_est28: *
29: em_est29: *
30: em_est30: *
31: em_est31: *
32: em_est32: *
33: em_est33: *
34: em_est34: *
35: em_est35: *
36: em_est36: *
37: em_est37: *
38: em_est38: *
39: em_est39: *
40: em_est40: *
41: em_est41: *
42: em_est42: *
43: em_est43: *
44: em_est44: *
45: em_est45: *
46: em_est46: *
47: em_est47: *
48: em_est48: *
49: em_est49: *
50: em_est50: *
51: em_est51: *
52: em_est52: *
53: em_est53: *

54: em_est22: *
55: em_est23: *
56: em_est24: *
57: em_est25: *
58: em_est26: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	347.4	7.2	473	22	R51152	R51152 y971e07.r1
2	339	7.0	431	21	T74128	T74128 y682a11.r1
3	299.2	6.2	559	23	H16894	H16894 ym39a03.r1
4	290	6.0	291	21	F12385	F12385 HSC39C111.n
5	263.8	5.5	517	43	A1176907	A1176907 EST320513
6	239.8	5.0	505	46	A1410862	A1410862 EST339135
7	201.4	4.2	422	45	A1233991	A1233991 EST330679
8	200	4.1	400	48	A1602045	A1602045 UI-R-GO-u
9	194	4.0	265	29	AA184901	AA184901 mt93e12.r
10	152	3.1	497	41	A1072568	A1072568 UI-R-C2-n
11	131.8	2.7	864	45	A1324553	A1324553 mr92d06.y
12	130	2.7	657	43	A1173222	A1173222 ud63a12.x
13	121.6	2.5	423	40	AA984429	AA984429 am86c08.s
14	121.4	2.5	633	30	AA208846	AA208846 mw75a01.r
15	120.8	2.5	1006	43	A1225600	A1225600 u106c05.y
16	120.6	2.5	613	34	AA512671	AA512671 v939g02.r
17	117	2.4	456	45	AA817854	AA817854 UI-R-A0-a
18	114.8	2.4	785	43	A1181867	A1181867 ud63a12.y
19	110	2.3	518	39	AA839539	AA839539 uc99a11.r
20	105.6	2.2	521	39	AA865338	AA865338 cg88d10.s
21	104.8	2.2	864	48	A1573566	A1573566 ud45f11.y
22	102.6	2.1	365	26	W45868	W45868 mc81h11.r1
23	101.8	2.1	625	47	A1526753	A1526753 u142e03.y
24	100.4	2.1	859	45	A1326571	A1326571 mm20g07.y
25	98.6	2.0	629	50	A1683839	A1683839 fx77g04.x
26	97.2	2.0	427	38	AA771447	AA771447 v45b04.x
27	96	2.0	567	39	AA873581	AA873581 oh68a12.s
28	94.8	2.0	462	29	AA166973	AA166973 z087e05.r
29	94.2	1.9	206	48	A1547643	A1547643 UI-R-C3-s
30	93.8	1.9	477	34	AA461838	AA461838 v195d04.r
31	93.8	1.9	315	48	A1611296	A1611296 tul6h02.x
32	92.6	1.9	667	28	AA106138	AA106138 mm20g07.r
33	91.8	1.9	586	26	W83411	W83411 mf28d06.r1
34	91.4	1.9	689	43	A1195467	A1195467 u152c06.x
35	90.2	1.9	820	43	A1194959	A1194959 u158h03.x
36	90	1.9	372	21	T61867	T61867 yb93a02.s1
37	87	1.8	360	20	T28497	T28497 EST746762.Hu
38	87	1.8	800	42	A1114892	A1114892 u149f07.x
39	86.6	1.8	530	43	A1195667	A1195667 u152c06.y
40	86	1.8	547	30	AA238246	AA238246 mx76d03.r
41	86	1.8	685	49	A1646857	A1646857 ud66f09.x
42	85.8	1.8	489	30	AA270973	AA270973 va84c09.r
43	85	1.8	475	44	A1265337	A1265337 uk01h06.y
44	84.8	1.8	392	30	AA221925	AA221925 mw13e10.x
45	84.6	1.8	666	44	A1323769	A1323769 mm20g07.x

ALIGNMENTS

RESULT 1
LOCUS R51152 473 bp mRNA
DEFINITION y971e07.r1 Soares infant brain IN1B Homo sapiens cDNA clone
IMAGE:38551 5' similar to gb:M4113 CONGULATION FACTOR VIII
PRECUSOR (HUMAN);, mRNA sequence.
ACCESSION R51152
NID 9813054

```

VERSION      RS1152.1  GI:813054
KEYWORDS
SOURCE       Homo sapiens
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
              Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 473)
AUTHORS      Haller,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
              Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
              Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
              Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and
              Wilson,R.
TITLE        The WashU-Merck EST Project
JOURNAL      Unpublished (1995)
COMMENT      Contact: Wilson RK
              Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: est@watson.wustl.edu
              High quality sequence stops: 351
              Source: IMAGE Consortium, LNLN.
              This clone is available royalty-free through LNLN; contact the
              IMAGE Consortium (info@image.lnl.gov) for further information.
              Seq primer: M13RPI
              High quality sequence stop: 351.
              Location/Qualifiers
                1..473
                /organism="Homo sapiens"
                /db_xref="GDB:411092"
                /db_xref="taxon:9606"
                /clone="IMAGE:38551"
                /clone_lid="Soares infant brain INIB"
                /sex="female"
                /dev_stage="73 days post natal"
                /lab_host="DH10B (ampicillin resistant)"
                /note="Organ: whole brain; Vector: Lactimid BA; Site.1: Not
                I; Site.2: Hind III; 1st strand cDNA was primed with a Not
                I - oligo(dT) primer [5',
                AACTGGAGAATTCGCGCGCCGAGCAATTTTCTTTTCTTTT 3'];
                double-stranded cDNA was ligated to Hind III adaptors
                (Pharmacia), digested with Not I and directionally cloned
                into the Not I and Hind III sites of the Lactimid BA vector.
                Library went through one round of normalization. Library
                constructed by Bento Soares and M.Fatima Bonaldo."
COUNT      116 a 118 c 104 g 132 t
IN
Query Match 7.2%; Score 347.4; DB 22; Length 473;
Best Local Similarity 96.8%; Pred. No. 1.3e-88;
Matches 419; Conservative 0; Mismatches 6; Indels 8; Gaps 6;

QY 4393 AGATTGGCCCTGAGGATGAGGTTCTGGGCTGGAGGACAGAGACCTACTAGAGGTGGC 4452
Db 1 AGATTGGCCCTGAGGATGAGGTTCTGGGCTGGAGG-ACAGGACCTACTAGAGGTGGC 59

QY 4453 CACTGACAGACCTGCGCACTGCGTACCTCTCCCTCCTAGCTCCAGGCACTGTCCTC 4512
Db 60 CACTGACAGACCTGCGCACTGCGTACCTCTCCCTCCTAGCTCCAGGCACTGTCCTC 119

QY 4513 CCTGGCTTGCCTTCTACCTTTGTCTAAATCTCTAGACAGACTGCTTGAAGCCTCTGA 4572
Db 120 CCTGGCTTGC-CTTCTACCTTTGTCTAAATCTCTAGACAGACTGCTTGAAGCCTCTGA 178

QY 4573 ATTACTATNCACTGCTGCTGCTTTCTTTGGGGGGGCGCAGAGGTCATCCAAATTTA 4632
Db 179 ATTACTATNCACTGCTGCTGCTTTCTTTGGGGGGGCGCAGAGGTCATCCAAATTTA 238

QY 4633 ACTTACTCTTACTATTTTCTGACAGTGCCTCCAGATTACTCTTCTTCCATATAAC 4692
Db 239 ACTTACTCTTACTATTTTCTGACAGTGCCTCCAGATTACTCTTCTTCCATATAAC 298

```

```

QY 4693 TAGCAAAAAGAGTGAAGAGAAACCT-GCATGAACCATTTCTCCCTGAAGTNGCC 4751
Db 299 TAGCAAAAAGAGTGAAGAGAAACCTGCATGAACCATTTCTCCCTGAAGTNGCC 358

QY 4752 CTCACAGTCACACAC-TCCTGTGTGTA-GAAAACTATGTS--ATGAACCTTGAA 4806
Db 359 CTCACAGTCACACACCTTCTCTGTGTGTAGAAAACTATGTGATGGAACCTTGAA 418

QY 4807 AAGATATTTATG 4819
Db 419 AAGATATTTATG 431

RESULT      2
LOCUS       T74128
DEFINITION  Y682a11.r1 Soares infant brain INIB Homo sapiens cDNA clone
IMAGE:22304.5', similar to gb:M14113 COAGULATION FACTOR VIII
PRECEDSOR (HUMAN);, mRNA sequence.
ACCESSION   T74128
MID         9690803
VERSION     T74128.1  GI:690803
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 431)
AUTHORS     Haller,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and
            Wilson,R.
TITLE       The WashU-Merck EST Project
JOURNAL     Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 4827
            High quality sequence stops: 373 Source: IMAGE Consortium, LNLN. This
            clone is available royalty-free through LNLN; contact the IMAGE
            Consortium (info@image.lnl.gov) for further information.
            Insert Length: 4827 Std Error: 0.00
            Seq primer: M13RPI
            High quality sequence stop: 373.
            Location/Qualifiers
              1..431
              /organism="Homo sapiens"
              /db_xref="GDB:394651"
              /db_xref="taxon:9606"
              /clone="IMAGE:22304"
              /clone_lid="Soares infant brain INIB"
              /sex="female"
              /dev_stage="73 days post natal"
              /lab_host="DH10B (ampicillin resistant)"
              /note="Organ: whole brain; Vector: Lactimid BA; Site.1: Not
              I; Site.2: Hind III; 1st strand cDNA was primed with a Not
              I - oligo(dT) primer [5',
              AACTGGAGAATTCGCGCGCCGAGCAATTTTCTTTTCTTTT 3'];
              double-stranded cDNA was ligated to Hind III adaptors
              (Pharmacia), digested with Not I and directionally cloned
              into the Not I and Hind III sites of the Lactimid BA vector.
              Library went through one round of normalization. Library
              constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT  110 a 110 c 84 g 120 t
ORIGIN

```


JOURNAL
MEDLINE
COMMENT
and its expression
C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
95277534

Contact: Genethon
Genethon Centre de recherche sur le Genome Humain
1, rue de l'Internationale, BP60 91002 Evry Cedex, FRANCE
Tel: 33169472800
Fax: 33160778698
Email: genexpress@genethon.fr
Single read.
Genexpress library: C: Genexpress_sequence_idt: ylc-39c11
Insert length: 1570 Std Error: 0.00
Seq primer: (-21)M13-universal
High quality sequence stop: 417.
Location/Qualifiers
1. .291

FEATURES

source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="21"
/clone_lib="normalized infant brain cDNA"
/sex="Female"
/tissue_type="total brain"
/dev_stage="3 months old"
/note="Organ: brain; Vector: lafmid BA; Site_1: HindIII;
Site_2: NotI; sex=Female; dev_stage=3 months old;
isolate-muscular atrophy patient; tissue-type-total
brain; total mRNA was oligo (dt) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the
lafmid BA vector. Clone library from B.Souares, Psychiatry
Dept. Columbia University, USA. Normalization method:
Bento Soares, P.N.A.S in press"

BASE COUNT 52 a 92 c 66 g 80 t 1 others
ORIGIN

Query Match 6.0%; Score 290; DB 21; Length 291;
Best Local Similarity 99.7%; Pred. No. 3e-72;
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

4391 CCAGATTGCCCTGAGAGGAGTCTGGGCTCGAGGACAGAGACCTACTAGAGGTG 4450
1 CCAGATTGCCCTGAGAGGAGTCTGGGCTCGAGGACAGAGACCTACTAGAGGTG 60
4451 GCCACTGCAGACCTGCCACTGCCCTCCTCCTCAGCTCCAGGGCAGTGTCC 4510
GCCACTGCAGACCTGCCACTGCCCTCCTCCTCCTCCTCAGGGCAGTGTCC 120
4511 TCCCTGGCTGGCTTACTTGTGCTTAATCTTACGACAGACCTCCTTGAAGCTCCT 4570
TCCCTGGCTGGCTTACTTGTGCTTAATCTTACGACAGACCTCCTTGAAGCTCCT 180
4571 GAATTACTATCATCATGCTCTGATTTCTTTGGTGGGGGCGAGGGGCAATTC 4630
GAATTACTATCATCATGCTCTGATTTCTTTGGTGGGGGCGAGGGGCAATTC 240
4631 TAATTACTCTTACTATTTCTTGCAGCTGCTCCAGATTACTCTCTCT 4681
TAATTACTCTTACTATTTCTTGCAGCTGCTCCAGATTACTCTCTCT 291

RESULT 5
A1176907/c 517 bp mRNA EST 20-JAN-1999
LOCUS A1176907
DEFINITION EST220513 Normalized rat ovary. Bento Soares Rattus sp. cDNA clone
ACCESSION A1176907
VERSION 93272545
KEYWORDS A1176907.1 GI:3727545
SOURCE Rattus sp.

ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 517)
AUTHORS Lee,N.H., Glodok,A., Chandra,I., Mason,T.M., Quackenbush,J.,
Kerlavage,A.R. and Adams,M.D.
TITLE Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat
Gene Index
JOURNAL Unpublished (1998)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2152025.

Contact: Lee, NH
ATCC
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@igf.org
Seq primer: M13-21.
Location/Qualifiers
1. .517

FEATURES

source
/organism="Rattus sp."
/db_xref="ATCC (inhost):2031594"
/db_xref="taxon:10118"
/clone_lib="ROVBY42"
/clone_lib="Normalized rat ovary, Bento Soares"
/note="Organ: ovary; Vector: pT73Pac; Site_1: EcoRI;
Site_2: NotI"

BASE COUNT 93 a 174 c 154 g 96 t
ORIGIN

Query Match 5.5%; Score 263.8; DB 43; Length 517;
Best Local Similarity 70.9%; Pred. No. 1.3e-64;
Matches 365; Conservative 0; Mismatches 147; Indels 3; Gaps 1;

2502 CAGAGCCCCGCGAGCTTCAAAAAGAACACGACATATTTATGCTGACGTGAGAG 2561
CAGGAGCCCCGCGGCTTACAGCAAAAGCGGCGACTATTCATCGCGCCGTGAAGAG 454
2562 CCTGGGATTATGGATGATGATGCTCCGACATGTTCAAGAAAGGGCTCAGAGTGC 2621
CTTGGGAGTACGGGAGTGTGTGCTGAGCTCCGGGACAGAGAGGGCTGAGTGG 394
2622 AGTGTCCCTCAGTTCAAGAAAGTTGTTCCAGAGATTACTGATGCTCTTACTCAG 2681
GACGAGCTAGGTTCCGAAAGTGTTCGTGATGACGAGAGGGTCTGACGAG 334
2682 CCTTATACCGTGGAGAACTAATTAACATTTGGAGCTCTGGGCGCATATTAAGACA 2741
CGCGTGCACCGTGGGAGACTGACGACACCTGGGGCTACTGGGCCATATACGGCG 274
2742 GAAGTGAAGATATATCATGTTACTTCAAGAAATCGAGGCTCGTCCATATTC 2801
GAAGTGAAGATATATCATGTTACTTCAAGAAATCGAGGCTCGTCCATATTC 214
2802 TATTCAGCTTATTTCTTATGAGAGATGAGAGAGAGAGAGAGAGAGAGAG 2861
TATTCAGCTTATTTCTTATGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAG 157
2862 TTGTCAGAGCTAATGAACCAAACTTACTTTTGGAAAGTGAACATATATGACACC 2921
TTGTCAGAGCTAATGAACCAAACTTACTTTTGGAAAGTGAACATATATGACACC 97
156 TTGTCAGAGCTAATGAACCAAACTTACTTTTGGAAAGTGAACATATATGACACC 97
2922 ACTAAGATGAGTTGAGTGCAGAAAGCTGGGCTTATTTCTGATGATGACCTGAGAAA 2981
ACGAGAGCGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 37
2982 GATGTCACACTAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3016
GATGTCACACTAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2

DB	243	ACGCACAGTCCGCGCATTTACAGGGCCACAGCGTCACCGGTGGCGGAGGGGGCGGACACACA	302
QY	3361	AAATGGCATGTACATCTCTATCCAGGTGTTTTTATGACAGTGGAAATGTTACATCCA	3420
Db	303	GGACGGGGGTGTGCACACCTGATCCCGGGGTGTTCCAAACCGGGAAGATGGCCCTCCA	362
QY	3421	AAGCTTGGAATTTGGCGGGTGGGAATGACCTTATTGGCGAGCATTCATACGTGGGTGAGC	3479
Lib	363	GAGCGGGGATCTGGCGGGTGGAGAGCGCTGTGGGCGACAGCACTACGCGGTGGGATAGC	421
RESULT	8		
	A1602045/c		
LOCUS	A1602045	400 bp	mRNA
DEFINITION	UI-R-GG-us-d-09-0-UI-s2	UI-R-GG	Rattus norvegicus CDNA clone
ACCESSION	A1602045		
	94611206		
VERSION	A1602045.1	GI:4611206	
KEYWORDS	EST.		
SOURCE	Norway rat.		
ORGANISM	Rattus norvegicus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
REFERENCE	1 (bases 1 to 400)		
AUTHORS	Bonaldo,M.F., Lennon,G. and Soares,M.B.		
TITLE	Normalization and subtraction: two approaches to facilitate gene		
	discovery		
JOURNAL	Genome Res. 6 (9), 791-806 (1996)		
MEDLINE	97044477		
COMMENT	On May 7, 1998 this sequence version replaced gi:3121068.		

CONTACT: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel.: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 Oligo-dT track not found, Not I site shown in beginning of sequence
 is likely internal to the message. cDNA library Preparation: M.B.
 Soares lab clone distribution: clones will be available through
 Research Genetics (www.reagen.com)
 Seq primer: M13 Forward,
 Location/Qualifiers
 1..400

```

/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-G0-us-d-09-0-UI"
/clone_lib="ui-R-G0"
/dec_stage="adult"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pT730-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI. The UI-R-G0
library is a normalized library constructed from a
mixture of rat tissues (nodose ganglia, dorsal root
ganglia, and trigeminal ganglia)."
The tag is a string of
6 nucleotides present between the Not I site and the
oligo-dT track. The library was constructed as described
by Bonaldo, Lennom and Soares, Genome Research 6:
791-806, 1996."

```

Query Match	4.1%	Score 200;	DB 48;	Length 400;
Best Local Similarity	70.5%	Pred. No. 2,28,46;		
Matches	282;	Conservative	0;	Mismatches 115; Indels 3; Gaps 1;

QY	2655	GAATTACTGANGGCTCTTACTACGCCCTTATACCGTGGAGAACTAATATAACATTG	2714
Db	399	GAGTTCCAGGACGACGCTTTCACGAGCGCGCGACCGTGGGAGACTGACCCACACCTG	340

QY	2715	GGACTCTCGGGGCATCTATTAAGAAGCAAGTTGAAGATAATATATGTTACTTTCAGA	2774
Db	339	GGGCTACTGGGGGCATCTTAAGGGGGGAAGTGAAGACAACATGTTGACTTCGGG	280
QY	2775	AATAGAGCTCTGTGCTCCCTATCTCTCTATTTCACTAGCTTATTTTCTTAAGAGAGATAG	2833
Db	279	AACCAAGGCTCTCGCCCATCATCTATTATCTCAGGCTGTGTCTCTTACCCGAAACCCAC	220
QY	2835	AGGCAAGAGAGACAGAACCTTAACAAAAAATTTGTTCACACCTTAATGAACCAAACTACTTT	2894
Db	219	GG---TGAGAGTGAACCAAGAGACAATTGTGTGGGCCCAAGACCAAAACTTACTTC	163
QY	2895	TGGAAGATGCAACATCTATTATGCAACCCACTAAAGATGATTTGACTGCAAACTCTGGCT	2955
Db	162	TGGAGAGTCCGGCCCCACATAGGCGCCACGAGAGGGGAATTTGACTGCAAGGCTGTGGCT	103
QY	2955	TATTTCTCTGATGTTGACCTGGAAAAAAGATGTGCATCTAGGCGCTGATTTGACCCCTTGTG	3014
Db	102	TATTTCTCCACGCGGACCTCGAGACAGACCTGCATCTCGGGGCTGATTTGGGCCACTGCTC	43
QY	3015	GTCGACACACTAACACACTGAACCCCTGCTCATTTGGAGAGC	3054
Db	42	GTGTGCACGCTTCACACTTGCACACCCCTGCAAGGCGCTC	3

RESULT 9
LOCUS AAI84901
DEFINITION
AAI84901 265 bp mRNA EST 17-FEB-1997
mt93612.r1 Soares mouse 3NDMS cdNA clone IMAGE:537486
5, similar to gb:105573 Mus domesticus coagulation factor VIII
mRNA, complete (MOUSE);, mRNA sequence.
AAI84901
NID g1768759
VERSION AAI84901.1 GI:1768759
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 265)
Marr,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubnque,T.
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Therling,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
On Sep 12, 1996 this sequence version replaced g1:1292355.

Contact: Mairra M/Mouse EST Project
 WashU-HHMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL ; contact the
 IMAG Consortium (info@image.llnl.gov) for further information.
 MGI:389478
 Seq primer: -28m13 rev2 from Amersham
 High quality sequence stop: 250.
 Location/Qualifiers
 1..265
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:637486"
 /clone_lib="Soares mouse 3MBMS"
 /sex="male"
 /tissue_type="Spleen"
 /dev_stage="4 weeks"
 /lab_host="Dh10B"

BASE COUNT

82 a 63 c 53 g 67 t

ORIGIN

/note=Vector: pT73-Pac (Pharmacia); with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATGTGAGTGAGGGAGCCGCCGCTGTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. RNA provided by Dr. Bertrand Jordan. Library went through three rounds of normalization, and was constructed by Bento Soares and M.Patrina Bonaldo."

Query Match	4.0%	Score 194	DB 29	Length 265
Best Local Similarity	86.3%	Prod. No. 9.6e-45		
Matches 226	Conservative 0	Mismatches 35	Indels 1	Gaps 1

4090 GGAAGTAAATGCTTGGAGACCTCAGGTGATATATCAAAAGAGTGGCTCAAGTGAATTC 4145

Db

61	GGACTTAATGCCCTGGCGACCTCAGGGAATGATCCAAAACAATGGTTGCAAGT-GACCTTAC	119
----	--	-----

Oy 4150 AGAGACAATGAAGTACACAGAGTAAGTACTCAGGAGTAATAATCTCGTTACCAGCA 4209
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 120 AAAAGACAATGAAGTACTCGAATTAATTAACCCAGGAGTGAAATCTCTCTTTAACAGCA 179

[illegible]

QY 4270 TTCAGATGCGCAAGTAAGGT 4291
| | | | | | | | | | | | | | |
Db 240 TATACAATGGCAAGTAAGGT 261

RESULT	10
AI072568	
LOCUS	AI072568 497 bp mRNA EST 11-FEB-1999
DEFINITION	UI-R-C2-n1-h-12-0-UI.s1 UI-R-C2 Rattus norvegicus cDNA clone
	UI-R-C2-n1-h-12-0-UI 3', mRNA sequence.

JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT On Nov 29, 1993 this sequence version replaced g1:635447.

```

SOURCE
1. .497
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/map="13"
/clone="UI-R-C2-n1-h-12-0-UI"
/clone_11b="UI-R-C2"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pC7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-C2
library is a subtracted library derived from the UI-R-C1
library, which is a subtracted library derived from the
UI-R-C0 library. The UI-R-C0 library consisted of a
mixture of individually tagged normalized libraries
constructed from rat placenta, adult lung, brain, liver,
kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day
embryo. The tag is a string of 3-5 nucleotides present
between the Not I site and the oligo-dT track which allows
identification of the library of origin of a clone within
the mixture. The subtracted library (UI-R-C2) was
constructed as follows: PCR amplified cDNA inserts from
UI-R-C1 clones from which 3' ESTs had been derived was
used as a driver in a hybridization with the UI-R-C1
library in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library) was
purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
UI-R-C2 library. This procedure has been previously
described (Bonaldi, Lennon and Soares, Genome Research 6:
791-806, 1996)"

```

Query Match	3.1%	Score 152;	DB 41;	Length 497;
Best Local Similarity	66.5%	Pred. No. 1.4e-32;		
Matches 218; Conservative	0;	Mismatches 110;	Indels 0;	Gaps 0;

DQ 1323 GTCCTGCCCCCGATGACAGAACTTTATAAAGTCAATATTTAACAATGGCCCTCAGCG 1382
| | | | | | | | | | | | | | | |
Dd 3 GCCGCCACCCTAAGAAGCGGGAGCCTGCAGAACCGGTACTGGGCGTGCGCCCTGGCCGC 62

Dc
63 ATCCGCAACAAATATAAGAAATCAGGTTCGTGGCTACACGCACGCCACTTACACACG 122

Df 123 CGGAGGCCACCCCGGTAAAGCTGGGCTCTTAGGCCACTGCTCTTACGGGGAGGTCGA 182

DQ 1503 GACACAGCTGTGATTATATTTAAGATCAAGCAGACCATATAACTCTACCCCTCAC 1502

Do 183 GACTCAGTGTGATTGTGTTCAGAACCGACGGACCGGCATACAACATCCACCCTCAT 242

QY 1563 GGAAATCACTGATGTGCCGTCCTTTGTAATTCAGAGAGATTACCAAAAAGCTGTAAACAATTGG 1622

Db 243 GGAGATCCGTGACGTTGGCGCTGTGCATCAGAGAGCGTCCACAGAGGGGTGAAGCACGTG 302

Qy	1623	AAGGATTTTCCAA	TCTGCCAGAGAAA	1650
Db	303	AAGGATTTGCCGAT	CCGCCACGAGAGA	330

RESULT 11
AI324553

	DEFINITION	
m92ad06.y1	Stratagene mouse embryonic carcinoma (#937317)	BUS
musculus cdna IMAGE:604907.5	similar to gb:M83837	MUS
fat globule membrane protein E8 mRNA,	complete (MUSEE);	MUS
	mRNA	

and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTCTGCTCTTAAGACGCG and 3' end primer CGACCTGCAGCTCGACACA.

BASE COUNT 300 a 198 c 224 g 274 t 10 others
ORIGIN

Query Match 2.5%; Score 120.8; DB 43; Length 1006;
Best Local Similarity 54.9%; Pred. No. 1.7e-23;
Matches 277; Conservative 0; Mismatches 222; Indels 6; Gaps 2;

```
QY 264 TACAAAAAGACTGTTTGTAGATTTCAGGATCCCTTTCAACATGCTAAGCAAG 323
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 227 TATAGAGAGGCCCTTTATTTGAGTACACAGATGACACCTTAGTAGACTATGACAA 286
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    324 CCACCTGATGGTCTGCTAGTCTTACCATCCAGCTGAGGTTTATGATACAGTCTC 383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    287 CCAGCCCTGGCTAGGCTTTTAGGCCCTGTCTCAAAAGCTGAAGTTGAAGTTTAT 346
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 384 ATTACACTTAAGAACATGGCTTCCATCCTGTCTGATGCTGTGTGATCTTAC 443
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 347 GTTCACTTAAGAACCTTGGCTCTAGGATCTTACACTTTTCATGCACATGGGGTAACTAC 406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 444 TGGAAAGCTTCTGAGAGGCTGATATGATGATGACACCAAGTCAAAAGGAAAGAGAT 503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 407 ACCAAGAGATATGAGGAGCGCTTACCTGACACACACACTGATTTTCAAGGGCTGAT 466
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 504 GATTAAGTCTTCCCTGGTGAAGCCATACATATGTGCGAGGTCTGAAAGAGATGT 563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 467 GACAAAGTCTTCCCGGCAACAGTATGTATGTGTGCTGATGCC--ATGAGCCAAGT 523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 564 CCATGGCCCTCTGACCCACTGTGCTTACTACTCATATCTTTCTCATGTGACCTGTA 623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 524 CTTGAGAGGGAGACAGCANTTGTGACCAAGATTTACCACTCCATGTTGATGCTCA 583
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 624 AAGACTTGAATTCAGGCTCATTTGAGAGCCCTACTAGTATGAGAGAGGAGTCTGCC 683
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 584 NAGATATTGCATCAGACATAGACCTCTATCTGTANMAAAGTTCTCTATAN 643
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 684 AAGGAA--AGACACAGACTTGCACAAATTATCTTTTGTCTATTTGATGAA 740
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 644 TAGGAAAAAGAAAAATATTGACCAAGAAATNTACTAATGTCTCTGTGTGATGAA 703
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 741 GGGAAAGTGGGACCTCAGAAACAA 765
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 704 AATCTCAGTGTATCTTGGAGAA 728
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: August 18, 1999, 14:42:40
Job time: 789 sec

